Frago, S., Nicholls, R., Strickland, M., Hughes, J., Williams, C., Garner, L., Surakhy, M., Maclean, R., Rezgui, D. A. R., Prince, S., Zaccheo, O., Ebner, D., Sanegre, S., Yu, S., Buffa, F. M., Crump, M., & Hassan, A. B. (2016). Functional evolution of IGF2:IGF2R domain 11 binding generates novel structural interactions and a specific IGF2 antagonist. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(20), E2766-E2775. https://doi.org/10.1073/pnas.1513023113

Peer reviewed version

License (if available):
Unspecified

Link to published version (if available):
10.1073/pnas.1513023113

Functional evolution of IGF2:IGF2R domain 11 binding generates novel structural interactions and a specific IGF2 antagonist

Susana Frago[1],Ryan D. Nicholls[2], Madeleine Strickland[2], Jennifer Hughes[1], Christopher Williams[2], Lee Garner[1], Mirvat Surakhy[1], Rory Maclean[1],Dellel Rezgui[1], Stuart N. Prince[1], Oliver J. Zaccheo[1], Daniel Ebner[3], Sabina Sanegre[1],Sheng Yu[4], Francesca Buffa[4], Matthew P. Crump[2¶] and A. Bassim Hassan[1¶].

[1]Tumour Growth Control Group, Oxford Molecular Pathology Institute, Sir William Dunn School of Pathology, University of Oxford, Oxford, OX1 3RE, UK. [2]Department of Organic and Biological Chemistry, School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK. [3]Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, OX3 7FZ, UK. [4]Department of Oncology, Old Road Campus Research Building, University of Oxford, Oxford, OX3 7DQ, UK

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Of the 15 extra-cellular domains of the Mannose 6-phosphate/ Insulin-like growth factor-2 receptor (M6P/IGF2R), domain 11 evolved a binding site for IGF2 to negatively regulate ligand bioavailability and mammalian growth. Despite the highly evolved structural loops of the IGF2: domain 11 binding site, affinity enhancing AB loop mutations suggest binding is modifiable. Here, we ask to what extent IGF2: domain 11 affinity, and its specificity over IGF1, can be enhanced, and if so, what is the structural basis of the mechanistic and functional consequences? Domain 11 binding loop mutants were selected by yeast surface display combined with high-resolution structure based predictions, and validated by surface plasmon resonance. We discovered novel mutations in the ligand interacting surface binding loops (AB, CD, FG and HI). Five combined mutations increased rigidity of the AB loop confirmed by NMR. When added to three independently identified CD and FG loop mutations that reduced the k_{off} by 2-fold, these mutations resulted in an overall selective 100-fold affinity improvement. The structural basis of the evolved affinity was improved shape complementarity established by inter-loop (AB-CD) and intra-loop (FG-FG) side chain interactions. The high affinity of the combinatorial domain 11-Fc fusion proteins functioned as ligand soluble antagonists or traps that depleted pathological IGF2 isoforms from serum, and abrogated IGF2 dependent signaling *in vivo*. An evolved and re-engineered high specificity M6P/IGF2R domain 11 binding site for IGF2 may improve therapeutic targeting of the frequent IGF2 gain of function observed in human cancer.

growth factor receptor | protein evolution | insulin-like growth factor 2 | binding kinetics | biological therapy

Introduction

Functional evolution of proteins is largely considered to occur by chance, frequently because of unpredictable and specific events that confer a structure based change in function sufficient for subsequent selection or 'gain of fitness' (1). One such evolutionary biochemical example is the initial acquisition and subsequent gain of affinity between the insulin-like growth factor 2 ligand and a single domain of a non-signaling M6P/IGF2 receptor (domain 11). The structural and functional basis of this evolutionary path, that has occurred over 150 million years of mammalian evolution, has already been reported (2). The question addressed here is whether the IGF2: domain 11 interaction has reached an optimal state in the context of IGF2 activation of signaling receptors and in the ligand clearance function of M6P/IGF2R, and how far can we extend the binding interaction in terms of structural, biophysical and functional properties.

Functionally, and unlike products of other mammalian imprinted genes, domain 11 is unusual because it specifically evolved to bind to an evolutionary conserved IGF2 ligand with high affinity (3-5). After binding, clearance of extra-cellular IGF2 occurs by receptor internalisation followed by IGF2 degradation in the lysosomal compartment (3). This mechanism supports the parental conflict theory of imprinting (6), as loss of M6P/IGF2R function results in IGF2 ligand over-supply and embryonic and placental overgrowth (7-10). Moreover, IGF2: domain 11 binding underpins M6P/IGF2R function as a tumour suppressor, and is the basis for the application of domain 11 as a soluble IGF2 ligand antagonist (11-14). The relative affinity of IGF2 for the signaling receptors, IGF1R and isoform A of insulin receptor (IR-A), is slightly lower (2-20 nM) (15), than the affinity for full length M6P/IGF2R (1-2nM) (16). IGF binding proteins (IGFBPS) 1-6 can achieve higher affinity (between 0.01-20 nM), but bind IGF1 and IGF2 and are proteolytically labile (17, 18). IGF2 affinity for domain 11 may therefore have reached an optimum that balances both *in vivo* cellular function (i.e. IGF2 must dissociate in late endosomes) and selectivity for IGF2 over IGF1. Excess IGF2, as occurs in during embryonic development and in tumours, appears to overwhelm the normal clearance function of M6P/IGF2R, to promote growth via IGF1R/ IR-A signaling receptor activation. In this instance, alternative mechanisms may limit free IGF2, including increased supply of IGFBPs and release

Significance

During 150 million years of mammalian evolution, the membrane bound mannose 6-phosphate receptor evolved high-affinity binding loops for Insulin-like growth factor 2. It remains unknown whether this highly evolved ligand interaction is optimal, and whether it can be further evolved. We addressed this question using a mutation and selection strategy that incorporated surface display and protein structure. Multiple mutations of all the binding loops were identified, and improved affinity by 100-fold when combined, yet retained IGF2 specificity. Structurally, IGF2 surface interactions with binding loops were reshaped, indicating that binding site evolution could not be predicted. High IGF2 affinity binding domains could selectively inhibit IGF2 dependent cell signaling, and may be applied in therapeutic IGF2 targeting in cancer.

Reserved for Publication Footnotes

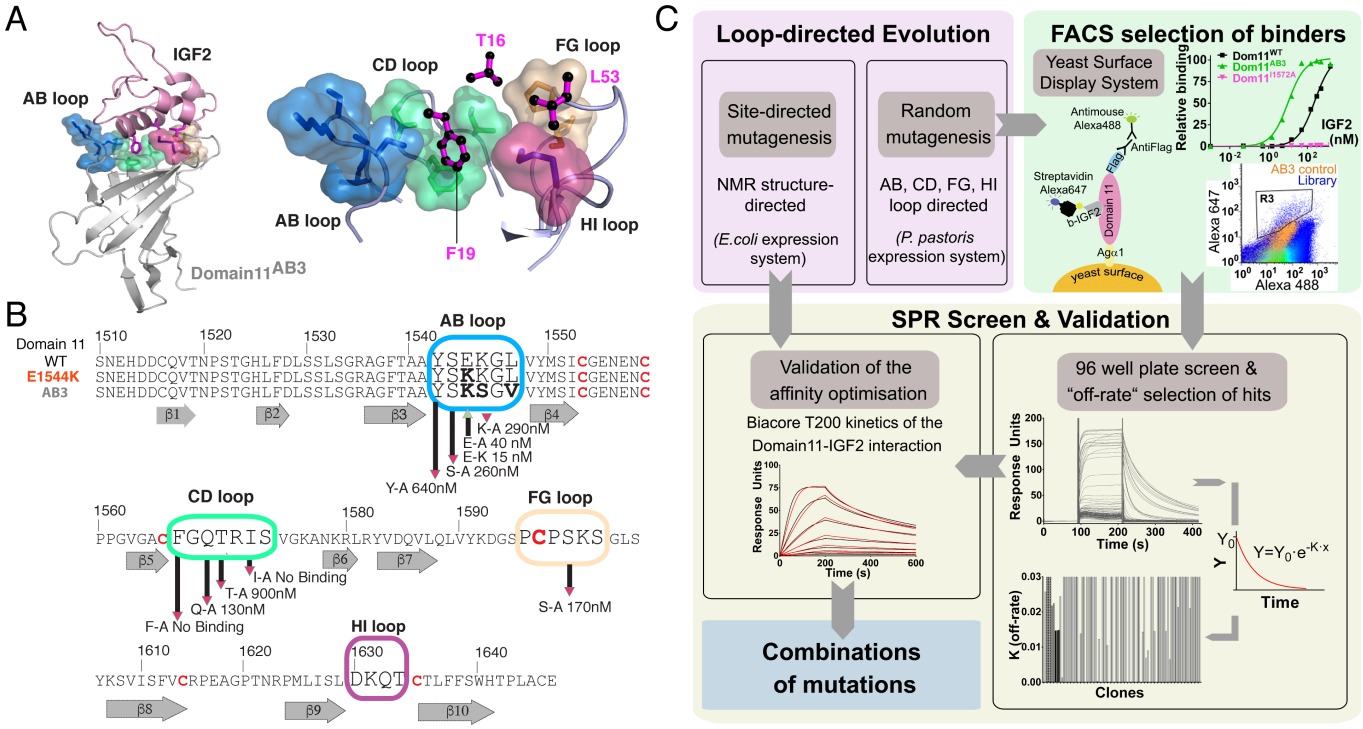


Fig. 1. Sequence and structure of the IGF2 binding site of domain 11 and *in vitro* forward evolution. (**A**) Left, ribbon representation of the solution structure of the complex between domain 11^{AB3} (grey) and IGF2 (pink). Surface representations are shown for the AB (blue), CD (green), FG (yellow) and HI (dark pink) loops. Right, close-up of the IGF2 binding site and interacting loops of domain 11^{AB3} and the main three-pronged interaction involving T16, F19 and L53 of IGF2. (**B**)Sequence of domain 11 secondary structure and binding loops. Mutations of Domain 11^{WT} and their effects on affinity for IGF2 are shown below the sequences. (**C**) Strategy for forward evolution. Mutant library generated by combining random mutagenesis in *P. pastoris* yeast surface display with flow cytometry screen (green) after validation of biotin-IGF2 binding to flag tag domain 11 protein on the yeast surface. Mutant clones are selected using a high throughput SPR screen, and with NMR informed site directed mutagenesis (purple), further validated (yellow) for IGF2 sensitivity and specificity using high sensitivity SPR of purified recombinant protein.

of soluble M6P/IGF2R that may trap the free ligand (13, 19). Full exploitation of selective IGF2 targeting in a wide range of cancers, using either antibodies or soluble receptors, remains an unmet need because of the lack of stable, high affinity and specificity therapeutic agents (20).

The IGF2 binding site of domain 11 in humans consists of a hydrophobic pocket centered on the CD loop, surrounded by polar and charged residues in the AB, FG and HI loops that complement surface charge on IGF2 (Fig. 1A) (2). The exception to this rule is an otherwise unfavorable charge-charge interaction between E1544 on the AB loop and D23 on IGF2 (4, 21, 22). Domain 13 indirectly interacts with the AB loop of domain 11, breaking this interaction in the multiple-domain receptor, and contributes to the stability of the complex by decreasing the 'off-rate' (k_{off}) of the IGF2 interaction (21). Mutations in the AB loop replace the effect of domain 13, increase the affinity for IGF2 (2, 22) and facilitated subsequent determination of a solution structure of the 24.2 kDa complex of IGF2 with domain 11^{AB3}, that incorporated three mutations of the AB loop (also known as AB3 or E1544K, K1545S, L1547V, K_D = 15 nM) (2). The structural basis of the interaction of IGF2: domain 11^{AB3} occurs with a relatively fixed conformation of the CD loop upon complex formation. The AB loop moves to accommodate IGF2 helix 1 that bears the critical recognition residues T16 and F19 (Fig. 1A, B), and the FG loop repositions between helices 2 and 3 of IGF2 to accommodate burial of a third important IGF2 residue, L53 (16, 21). The final complex forms a series of complementary hydrophobic surfaces and supports a range of H-bonding and salt bridging interactions (2). Although the IGF1 ligand is highly homologous to IGF2, particularly those residues involved with binding IGFBPs and signaling receptors, T16 of IGF2 confers specificity over IGF1 (where it is replaced by an alanine) for domain 11, most likely through packing interactions that precisely define the spacing of F19 and L53 (2). NMR structures of domain 11 from opossum, echidna, zebrafish and chicken all share similar β-barrel topology, yet IGF2 does not bind to domain 11 in birds and fish. Binding co-evolved with mammalian evolution, as primitive mammals (monotremes) bind IGF2 with ten-fold lower affinity (K_D = 250-400 nM *vs* human domain 11^{WT} K_D = 40-60 nM) (2, 23-25). The acquisition of the IGF2: CD loop hydrophobic interaction thus appears the main structural change during initial binding evolution, a chance event attributed to acquisition of an exon splicing enhancer that altered the amino acid coding of the CD loop (2). As mammalian domain 11 evolved, the surface surrounding the IGF2 binding region provided by AB, FG and HI loops displayed a loss of charged residues, increased volume and hydrophobicity, combined with the correct shape complementarity (2).

In order to address to what extent human IGF2: domain 11 affinity can be further evolved, we evaluated a multi-step approach to forward evolution. We developed *in vitro* directed evolution with yeast cell surface display, combined with structural directed mutagenesis and surface plasmon resonance affinity selection. The directed evolution of proteins to generate adaptive and positive selected functions is a powerful strategy, but is known to be dependent on a number of essential factors. These include the design and selectivity of the screening strategy, the epistatic effects of permissive existing and co-existing mutations on protein stability, and the practical limitations of generating comprehensive library diversity in bacteria (26-28). We identified novel and independent gain of binding function mutations of all the binding loops of domain 11, yet with retained binding specificity for

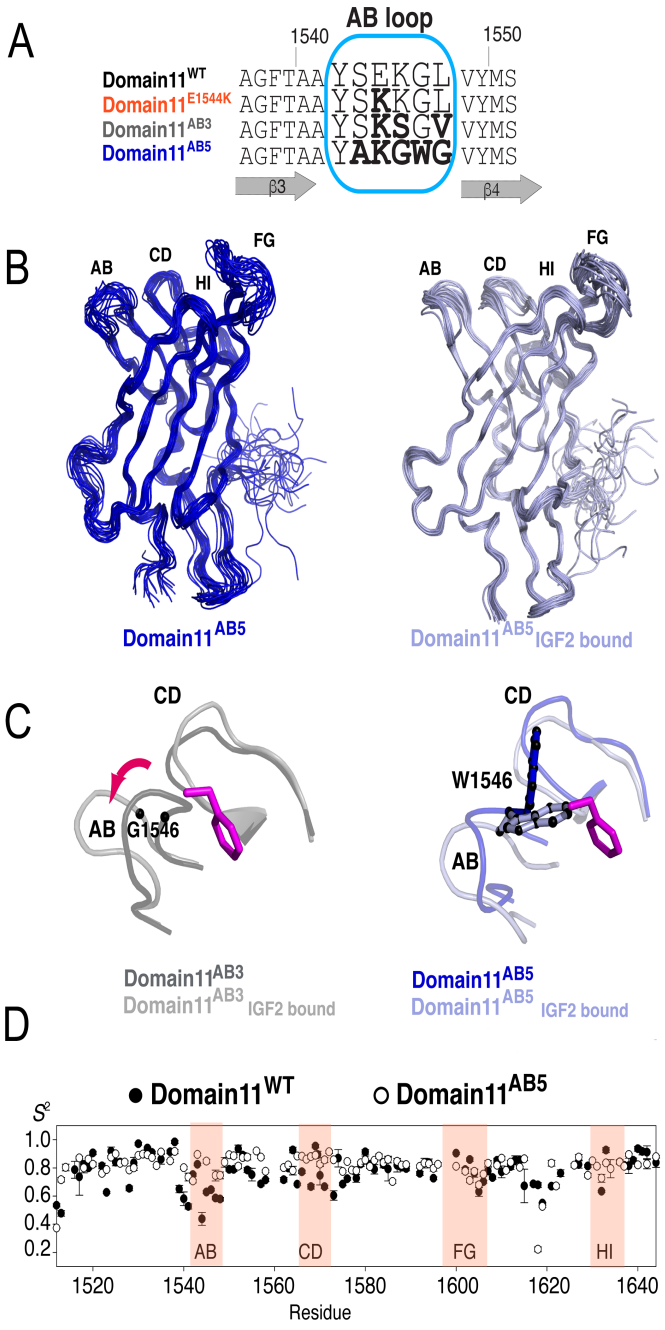


Fig. 2. Comparison of domain 11^{WT}, domain 11^{AB3} and domain 11^{AB5} high-resolution structures. (**A**) Sequence comparison of the AB loop from domain 11^{WT} and mutants. (**B**) High-resolution NMR structural ensembles of the lowest twenty energy models for free domain11^{AB5} (dark blue, PDB:2M6T) and bound to IGF2 (light blue, PDB:2M68). Structural statistics are given in in SI Appendix-Supplementary Table 4. (**C**) Key residues of the free forms (dark grey) of domain 11^{AB3} (left; PDB:2L2A) and domain 11^{AB5} (dark blue, right) and corresponding IGF2 bound forms of domain 11^{AB3} (light gray, PDB:2L29) and domain 11^{AB5} (light blue). In domain 11^{AB5} W1546 closes into IGF2 F19 (pink, modelled from 2L29) whereas in domain 11^{AB3} G1546 (highlighted by red arrow) moves outwards and indicates the large conformational change of the AB loop upon complexation. (**D**) Lipari-Szabo model-free S^2 parameters showing for domain 11^{WT} flexible AB, BC and CD loops, shown by S^2 values below 0.7. These loop regions are more rigid in domain 11^{AB5}, consistent with NOE, R_1 and R_2 data (SI Appendix-Supplementary Fig. 1-4).

IGF2. By also introducing structural informed mutations to those selected in the first round of directed evolution of ligand interacting residues, we iteratively combined individually identified loop mutants to generate domain 11 variants with even higher additive affinities, slower dissociation rates and improved thermodynamic properties. These novel high affinity mutant combinations generate novel structural mechanisms of ligand-receptor binding site interactions, and are the basis of specific soluble ligand traps for IGF2 that antagonise IGF2 function *in vivo*.

Results

Domain 11 yeast surface display and selection of high affinity AB loop mutants

For the directed evolution screen, we combined domain 11 surface plasmon resonance binding kinetics (SPR) (22), with an Agα1 yeast surface display system in *P. pastoris* (29). Mutants of an N-terminal flag-domain 11 were expressed, and cell surface expression quantified with an anti-flag antibody and Alexa488 labeled secondary antibody (Fig. 1C). Following binding of biotinylated human mature IGF2$^{1\text{-}67}$ to live yeast that express domain 11 controls, affinity for domain 11^{WT}, domain 11^{AB3} and a non-binding mutant domain 11^{I1572A} were quantified with Alexa647 labeled streptavidin (Fig. 1C, top right). In the first round of selection, yeast expressing mutant domain 11s were flow-sorted on binding to a low concentration of IGF2$^{1\text{-}67}$ (5 nM). Gates were based on the binding to domain11^{WT} and domain11^{I1572A} (non-binding mutant) and thresholds based on domain 11^{E1544K} or domain 11^{AB3} (higher IGF2$^{1\text{-}67}$ affinity, Fig. 1C) (22). Following culture of sorted clones, mutants were sub-cloned into a soluble protein expression vector, transfected cells plated and single clones screened in a 96 well format using a Biacore 3000 with k_{off} ($y=Y_0.e^{-k_{off}X}$) determination (Fig. 1C). Yeast expressing soluble mutant his$_6$-domain 11s with slow k_{off} were expanded, protein purified and SPR validation performed using a more sensitive BIAcore T200. The screen was applied to mutagenesis of the AB loop (^{1542}YSEKGL^{1547}) where the E1544K gain of function mutation was incorporated (^{1542}YSKKGL^{1547}) (SI Appendix-Supplementary Table 1). We identified an AB loop with five mutations (domain 11^{AB5}) and with a ten-fold higher IGF2$^{1\text{-}67}$ affinity compared to domain 11^{WT} (K_D 5.07 nM *vs* 46-64 nM, SI Appendix-Supplementary Table 2), and no detectable binding to IGF1. The five AB loop mutations (^{1542}YAKGWG^{1547}) in domain 11^{AB5} generated a very different combination of amino acids than those identified in domain 11^{AB3} (^{1542}YSKSGV^{1547}) (Fig. 2A).

In parallel, *P. pastoris* cultures of non-mutagenised domain 11^{AB3} controls were screened for IGF2$^{1\text{-}67}$ binding, and a surprising number of clones were obtained with slow k_{off}. These selected clones arose as a consequence of random mutation during bulk culture. No mutations were selected in the β-barrel structure, but two spontaneous loop mutants were identified that bound IGF2$^{1\text{-}67}$ with high affinity (SI Appendix-Supplementary Table 3). The first was an FG loop mutant (P1597H), with a four-fold reduction in k_{off} associated with a ten-fold improvement in affinity compared to domain 11^{AB3} (K_D = 1.23 nM vs 15.3 nM, respectively). The second mutant was in the hydrophobic CD loop (Q1569R) with enhanced affinity (K_D= 2.30 nM). Although mutations within domain 11 that altered IGF2 affinity were expected, a mutant of the CD loop was unexpected because of the importance of hydrophobicity and shape complementarity during evolution of the binding site (2), and because Q1569A had little impact on IGF2 affinity (22).

Structure of the domain 11^{AB5} mutant informs selection of non-AB loop mutants

To determine how the newly identified AB loop mutations in domain 11^{AB5} altered the IGF2 binding site, high-resolution NMR structures of domain 11^{AB5} were determined in both the

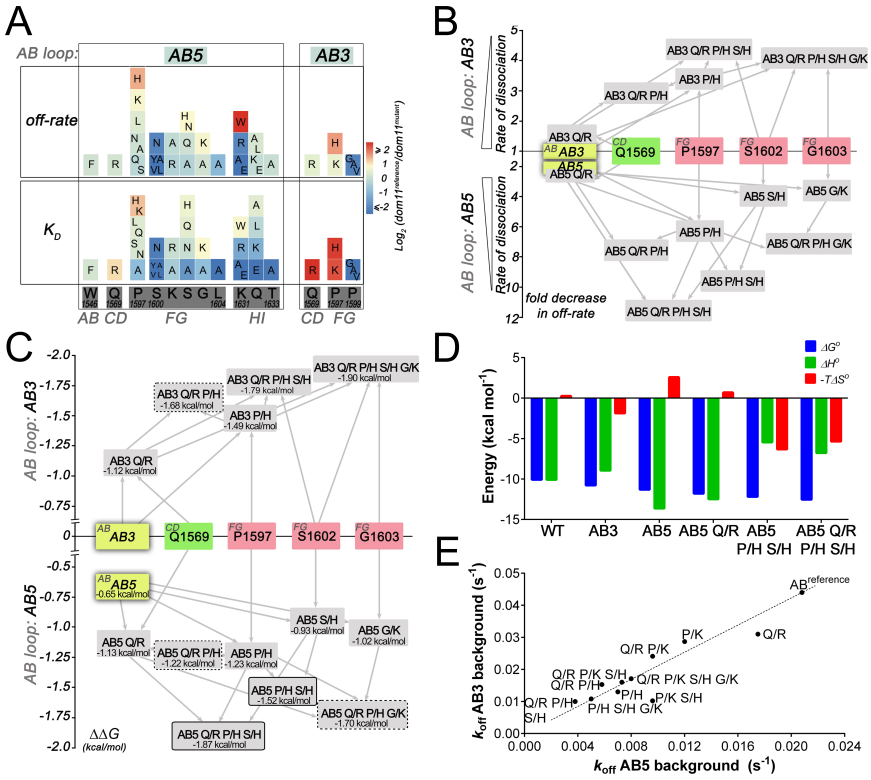


Fig. 3. The effects of single and combined mutations of the IGF2 binding site in human domain 11^{AB3} and domain 11^{AB5}. (A) Heat map of IGF2binding to single loop point mutants on domain 11^{AB5} (left) and domain 11^{AB3} (right) backgrounds.Top panel, k_{off}, Bottom panel,K_D. Scale; $\log_2$ of the domain 11^{reference}/ domain 11^{mutant} ratio, where domain 11^{reference} is either domain 11^{AB5} or domain 11^{AB3}. Reference value (green), increase of the affinity (red, decrease in K_D or k_{off}), decrease in affinity (blue, increase in K_D ork_{off}). **(B)** Effect of combinations of loop mutants on IGF2 k_{off} compared to domain 11^{AB3}. **(C)**Effect of combined loop mutants on IGF2 binding free energy (ΔG°) compared to domain 11^{AB3}. ΔΔG° calculated as ΔΔG° = ΔG°domain 11^{mutant} − ΔG°domain 11^{AB3} = RTln ($K_D$domain 11^{mutant}/$K_D$domain 11^{AB3}). Additive and non-additive contributions of single point mutations are shown with a box solid line and dashed line, respectively. Raw data in SI Appendix-Supplementary Tables 2, 5 and 6. **(D)** Thermodynamic profiles of the IGF2 binding of domain 11 mutants shown. Data fitting of the temperature dependence of the dissociation constant according to the van't Hoff equation. **(E)** Comparison of k_{off} between domain 11^{AB3} and domain 11^{AB5} with combined mutations.

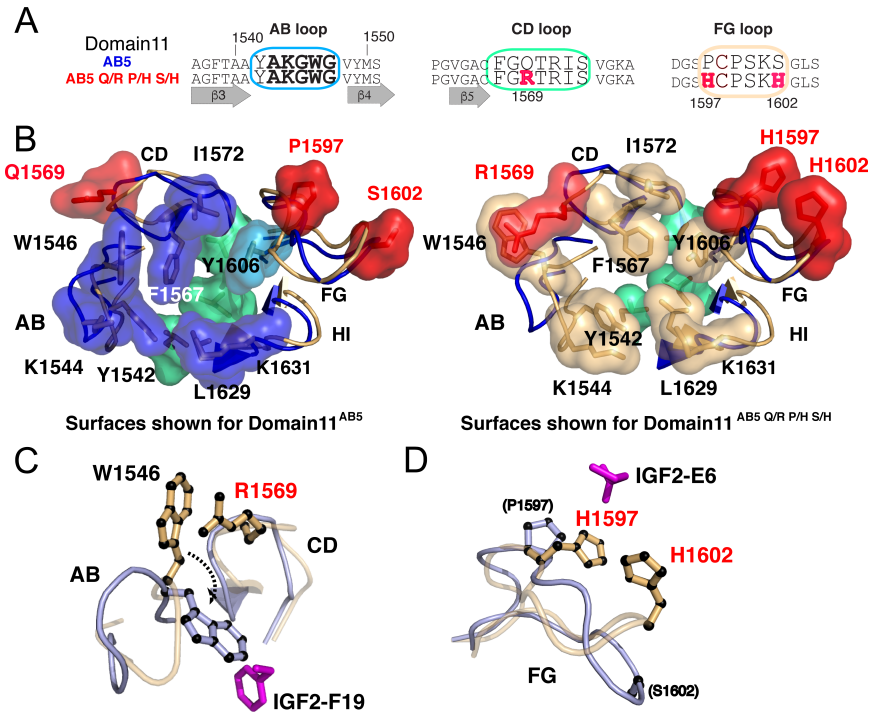


Fig. 4. Comparison of domain 11^{AB5} and domain 11$^{AB5\text{-}Q1569R\ P1597H\ S1602H}$ high-resolution structures. (A) Sequence comparison of the AB, CD and FG loop from domain domain 11$^{AB5\text{-}Q1569R\ P1597H\ S1602H}$ and domain11^{AB5}. **(B)** Superposition of free domain 11^{AB5} (blue loops) and domain 11$^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$ (gold loops) with residue surfaces shown for domain 11^{AB5} (left) and domain 11$^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$ (right). Hydrophobes within or surrounding the binding site are shown in blue (domain 11^{AB5}) or gold (domain 11$^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$) with underlying foundation residues in green (V1574, L1629 and L1636). The mutations Q/R 1569, P/H 1597 and S/H 1602 surfaces are shaded red. **(C)** Comparison of the AB and CD loop orientation and **(D)** the FG loop between the free form of domain 11$^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$ (orange) and domain11^{AB5} in the IGF2 bound form (light blue). The re-orientation of W1546 required for domain 11$^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$ to achieve the same bound conformation, is highlighted by the dashed arrow in **(C)**.

free and IGF2$^{1\text{-}67}$ bound forms (Fig. 2B) (2, 30). Domain 11^{AB5} retained the characteristic domain 11 flattened β-barrel fold (SI Appendix-Supplementary Table 4 and 5) that conserves the primary IGF2$^{1\text{-}67}$ binding site (Fig. 2A, SI Appendix-Supplementary Fig. 5A) (2). However, the AB loop adopted a significantly different structure in domain 11^{AB5}. The orientation of the side-chain of W1546, which substitutes for G1546, was not well-defined in the ensemble of structures, but appears to re-orient adjacent aromatic residues Y1542 and F1567, displacing the AB loop away from the IGF2 binding region. In the domain 11^{AB3} structure, this portion of the AB loop is hinged around the small G1546 residue and upon complex formation, the AB loop moves to accommodate IGF2$^{1\text{-}67}$ (Fig. 2C, 2D) (2). In the free form of

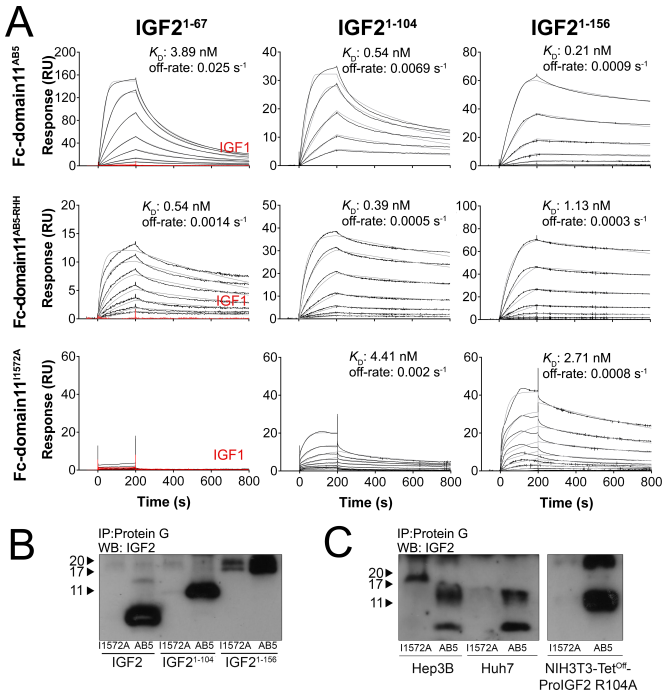


Fig. 5. . **Fc-domain 11$^{AB5/\ AB5\text{-}RHH}$ binds IGF2 isoforms with high affinity and selectivity**. (**A**) Surface plasmon resonance sensorgrams of the interaction between Fc-domain 11^{AB5}, Fc-domain 11$^{AB5\text{-}RHH}$ and Fc-domain 11^{I1572A} (controlnon-IGF2 binding) with IGF1 (red), IGF2$^{1\text{-}67}$, IGF2$^{1\text{-}104}$ and IGF2$^{1\text{-}156}$. Recombinant IGF1 and the different IGF2 forms were injected at concentrations ranging from 64 nM to 0.25 nM over Fc-domain11^{AB5} immobilised on a CM5 surface by antibody capture. (**B**) Fc-domain11^{AB5} or Fc-domain11^{I1572A} pull-down assay of different recombinant IGF2 isoforms. (**C**) Fc-domain 11 pull-down assay as in (**B**) of the different IGF2 isoforms produced by tumour cell lines. Supernatants of the HCC cell lines Hep3B and Huh7, and of the NIH 3T3 control cell line expressing pro-IGF2^{R104A} were incubated with Fc-domain 11^{AB5} or with Fc-domain 11^{I1572A} as a control.

domain 11^{AB5}, the AB loop more closely resembles the bound form and no equivalent rearrangement is necessary (Fig. 2C, 2D). ^{15}N-relaxation data for unbound domain 11^{WT} and domain 11^{AB5} were recorded at two field strengths (600 and 900 MHz) and revealed an overall increase in rigidity in the AB loop of domain 11^{AB5} (SI Appendix-Supplementary Fig 1-4, Fig. 2E). The binding of IGF2$^{1\text{-}67}$ to domain 11^{AB5} does however draw the W1546 aromatic side-chain inwards to the binding site. It is now well-defined in the ensemble of NMR structures and together with Y1542 and F1567, these three aromatic residues interact with the primary binding epitope, F19, from IGF2$^{1\text{-}67}$ (2, 21). K1544 was observed to form a salt bridge with D23 of IGF2 in the complex with domain11^{AB3} and most likely fulfills a similar role here. The K1545S mutation in domain 11^{AB3} is replaced in domain 11^{AB5} by K1545G, and therefore there is loss of hydrogen bonding capability to Q18 of IGF2$^{1\text{-}67}$ from the serine OH. Thus, of the five mutations in the AB loop of domain 11^{AB5}, these data implicate the structural importance of the specific G1546W mutation for the gain in affinity (Fig. 2C, 2D).

The Robetta server (http://robetta.bakerlab.org/alascansubmit.jsp) was used to perform interface alanine scanning mutagenesis on the IGF2$^{1\text{-}67:}$ domain 11^{AB3} complex to identify additional 'hotspot' loop residues important for complex formation. Mutations of these residues in domain 11^{AB5} and SPR affinity are shown (SI Appendix-Supplementary Table 2 and Fig. 3A). In comparison with the AB loop, the FG loop undergoes a pronounced structural change upon complexation in domain 11^{AB5}, suggesting potential for further optimisation of this sequence beyond P1597H. Mutations of FG loop residues K1601R, S1602H, S1602N and G1603K all exhibited up to 2-fold improvement in affinity (k_{off}), suggesting that the AB and FG loops can function in combination. Introduction of structurally predicted site directed mutations in the HI loop only slightly improved affinity, although a number of single point mutations could be introduced that appeared to slow the k_{off} (e.g. Q1632A and K1631W, SI Appendix-Supplementary Table 4, Fig. 4A). Finally introduction of W1546F into the AB loop reduced the k_{off} by two fold, yet retained the similar overall K_D (4.35 nM) as a result of a reduced k_{on}.

Directed evolution of domain 11 HI loop ligand interacting residues selects a higher affinity mutant.

The affinity of IGF2 for domain 11^{AB5} was at least two fold higher than our highest previously detected mutant. As we had already discovered single point mutations of the FG and CD loops that enhanced domain 11^{AB5} affinity, we systematically determined the potential for the remaining interacting HI loop in domain 11^{AB5} to specifically stabilise the IGF2 binding interaction. We first introduced a library of mutants for four ligand proximal residues of the HI loop (^{1630}DKQT^{1633}) into domain 11^{WT}, so that we could discriminate clear affinity gains compared to domain 11^{AB5}, and screened >1x10^5 fully representative clones representing every possible amino acid combination (SI Appendix-Supplementary Table 1). After flow sorting and subcloning 2.2 x 10^5 mutants into the soluble expression system, over >1000 clones were screened by SPR (Fig. 1C, SI Appendix-Supplementary Table 3). Only one mutant amongst several false positives was identified (^{1630}HFQS^{1633}) with a slower k_{off} than domain 11^{WT}.

Combinations of domain 11 loop mutants improves affinity 100-fold with retained IGF2 specificity

The first round of directed evolution combined with structural predictions had evolved domain 11 with respect to higher IGF2$^{1\text{-}67}$ affinity *in vitro,* with novel and independent AB, CD, FG and HI loop mutants. Moreover, the specific mutations in the CD and FG loops on AB loop background of either of domain 11^{AB5} or domain 11^{AB3}, modified the k_{off} and overall affinity by at least two fold (Fig. 3A and SI Appendix-Supplementary Tables 2 and 3). We determined whether the individual loop mutants identified by either the screens or by structural predictions could be combined to generate even higher binding affinity, but with retained IGF2$^{1\text{-}67}$ specificity. In particular, CD loop (Q1596R^{screen}), FG loop (P1597H^{screen}/K^{structure}, S1602H^{structure}, G1603K^{structure}) and HI (K1631W^{structure}, ^{1630}HFQS$^{1633/\ screen}$) mutations were therefore selected as candidates. Successive additions of the majority of mutations resulted in additive at best, rather than synergistic affinity effects, with improved binding kinetics on both domain 11^{AB3} and domain 11^{AB5} backgrounds (Fig. 3B, 3C and SI Appendix-Supplementary Table 6). The addition of the HI loop mutants K1631W and ^{1630}HFQS^{1633} often failed to result in stable expressed protein. The introduction of multiple mutations (up to 11) into domain 11^{AB5} also appeared less tolerated in terms of protein yields, than similar mutations in domain 11^{AB3}, indicating potential AB loop dependent epistatic effects on protein folding and stability (SI Appendix-Supplementary Table 6).

Thermodynamic parameters reflected the changes in affinity for IGF2$^{1\text{-}67}$ between domain 11^{WT}, AB loop mutants (domain 11^{AB3} and domain 11^{AB5}), double loop mutants domain 11$^{AB5\text{-}Q1569R}$ (AB+CD loops) and domain 11$^{AB5\text{-}P1597H\ S1602H}$ (AB+FG loops), and the triple loop mutant domain 11$^{AB5\text{-}Q1569R\ P1597H\ S1602H}$ (AB+CD+FG loops, also referred to as domain 11$^{AB5\ Q/R\ P/H\ S/H}$, and later abbreviated as domain

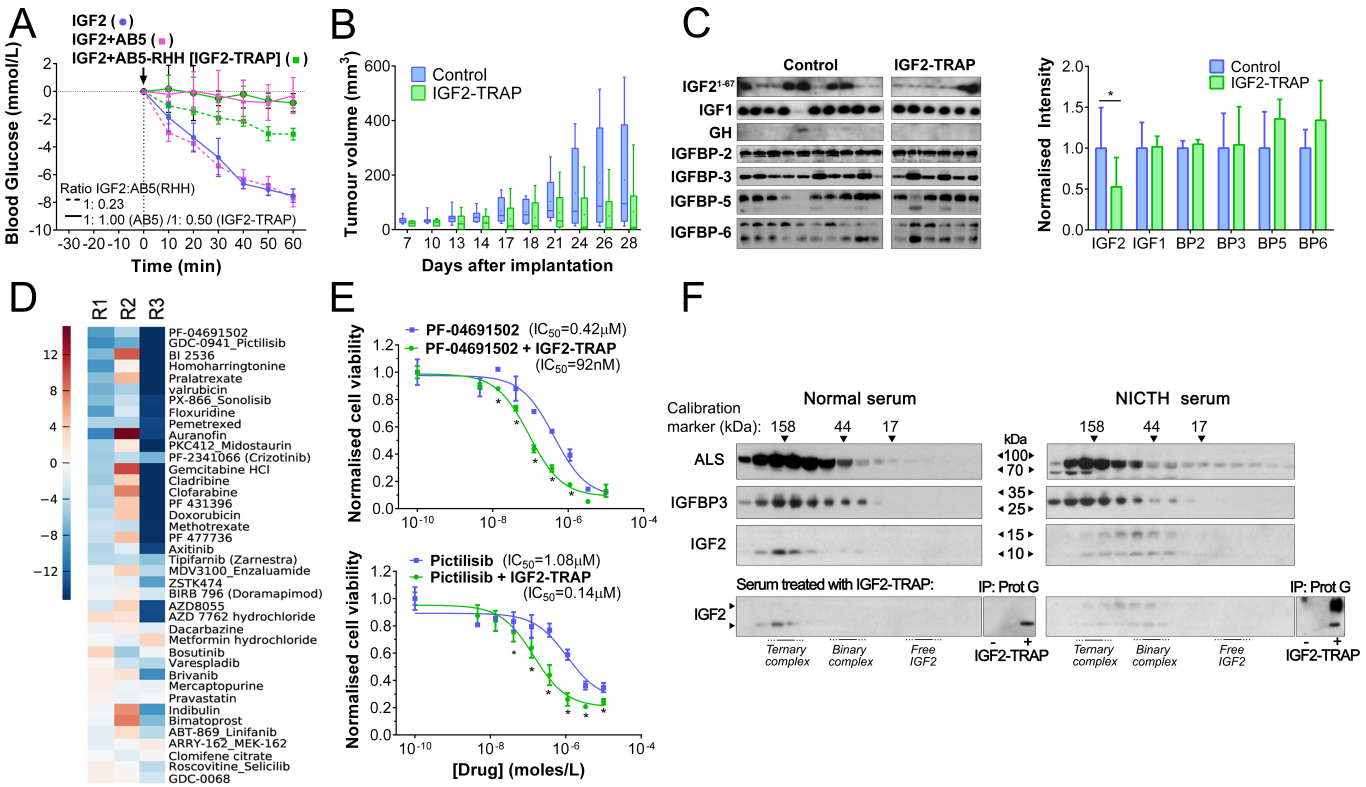


Fig. 6. Fc-domain 11^{AB5} and Fc-domain $11^{AB5\text{-}RHH}$ (IGF2-TRAP) inhibit IGF2 signaling *in vivo*. (**A**) Fc-domain 11^{AB5} and Fc-domain $11^{AB5\text{-}RHH}$ (IGF2-TRAP) abrogate an $IGF2^{1\text{-}67}$-induced hypoglycaemia in a mouse model. Mice were anaesthetized (time= -30 min) and blood glucose levels were allowed to stabilize for 30 mins (expressed relative to this blood glucose level). Subsequently (time= 0 min), they received 1 mg kg^{-1} $IGF2^{1\text{-}67}$ alone (n=4), or premixed with Fc-domain 11^{AB5} or with Fc-domain $11^{AB5\ RHH}$ at a molar ratio of 1:1 (n=3) (p=0.0133) or 1:0.5 (n=3) (p=0.0023), respectively. When using a molar ratio of 1:0.23, Fc-domain $11^{AB5\ RHH}$ is a more efficient IGF2 antagonist than Fc-domain 11^{AB5} (p=0.0026). Statistics 2 way ANOVA with Bonferonni post test. (**B**) IGF2-TRAP reduces IGF2 dependent xenograft growth (SKNMC-$IGF2^{1\text{-}67}$). $5x10^6$ cells per injection site in CD-1 nude mice, single infused concentration of IGF2-TRAP (40 mg kg^{-1} per week) (green, n=7, n=2 injection error, n=1 unexplained death) or PBS control (blue, n=10). (p=0.002, Wilcoxon test across all time points). (**C**) IGF2-TRAP administration resulted in reduced levels of serum IGF2 independent of IGF1, GH and IGFBPs levels (Day 28, control PBS n= 10 out of 10, IGF2-TRAP n=6 out of 7). (**D**) Oncology drug synergistic screenin SKNMC-FLAG-$IGF2^{104}$ with IGF2-TRAP. Heat maps show top-ranked drugs (30 shown, blue for synergy, red for antagonism) with low p-values (RP method) of the interaction score at 1 μM and the IGF2-TRAP (3 replicates R1-3). (**E**) Validation dose response curves for PI3kinase inhibitors (PF-04691502 and Pictilisib)in the presence (green) and absence (blue) of the IGF2-TRAP. IC_{50} values are shown. Leftward shift indicates synergism ($P < 0.0001$ when comparing the IC_{50} of drug alone *vs* drug + IGF2-TRAP, F-test). Asterisks indicate the concentrations at which there are synergistic interactions ($Q > 1.15$). (**F**) IGF2-TRAP modifies the molecular distribution of IGF2 in human serum. Normal (left) and NICTH (right) sera were fractionated in a gel filtration column at neutral pH, alongside a molecular weight calibration marker, before and after incubation and depletion with IGF2-TRAP-loaded protein G beads. Elution fractions were analysed by western blot.

$11^{AB5\text{-}RHH}$) (22, 31) (Fig. 3D, SI Appendix-Supplementary Fig. 6). The temperature dependence of the dissociation constant according to the van't Hoff equation showed the interaction of $IGF2^{1\text{-}67}$ with domain 11^{WT} appeared enthalpically driven ($\Delta H<0$), with a relatively small and un-favourable entropic contribution ($-T\Delta S>0$) (Fig. 3D). Mutations in the AB loop appeared to have different effects. Domain 11^{AB5} retained a similar unfavourable thermodynamic profile as domain 11^{WT}, with an unfavourable entropic contribution, whereas in domain 11^{AB3}, the entropic contribution appeared more favourable, as we had previously shown for the E1544K AB loop mutant (22). Addition of mutations to domain 11^{AB5} via CD or FG loops introduced more favourable changes in the entropic contribution and k_{off}, similar to domain 11^{AB3} (Fig. 3D, E). Further mutations in the CD and FG loops of domain 11^{AB5} introduced overall improvement in the free energy of binding (ΔG^o), with varied entropic and enthalpic compensation, consistent with the affinity improvement of the triple loop mutant domain $11^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$ (Fig. 3D, E).

Four higher affinity and specificity $IGF2^{1\text{-}67}$ binding versions of domain 11 were identified with at least a 100-fold improved affinity compared to domain 11^{WT} and 5-10-fold high affinity compared to the full length M6P/IGF2R. Within the domain 11^{AB5} background, domain $11^{AB5\text{-}Q1569R\ P1597H\ S1602H}$ (K_D= 0.65 nM pH 7.4/ = 0.71 nM pH 6.5) and domain $11^{AB5\text{-}Q1569R\ P1597H\ G1603K}$ (K_D= 0.87 nM pH 7.4/ = 0.69 nM pH 6.5) were identified, and were independent of pH between 6.5 and 7.4 (SI Appendix-Supplementary Table 6 and 7). For domain 11^{AB3} background, domain $11^{AB3\text{-}Q1569R\ P1597H\ S1602H}$ (K_D= 0.74 nM pH 7.4/ = 0.40 nM pH 6.5) and domain $11^{AB3\text{-}Q1569R\ P1597H\ S1602H\ G1603K}$ (K_D= 0.78 nM pH 7.4/ = 0.43 nM pH 6.5) were selected as the affinity (slower k_{off}) was observed to increase at lower pH.

X-ray crystallographic structure of domain $11^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$ that combines eight loop mutations

The X-ray crystal structure of one of the highest affinity domains, domain $11^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$, was solved to 2.8 Å resolution. This showed several unexpected structural rearrangements of the hydrophobic core compared to domain 11^{AB5} (SI Appendix-Supplementary Table 8). In domain11^{AB5}, W1546 adopted multiple conformations, but these were bounded by a neighbouring surface defined by Q1569 (Fig. 4A,B). This is

reversed in domain $11^{AB5\text{-}Q1569R\ P1597H\ S1602H}$ where R1569 lines the IGF2 binding site and engages W1546 in a cation-π interaction, directing the indole ring away from the interface (Fig. 4B). R1569 and W1546 are separated by 3.5 Å, consistent with the expected range for this type of interaction (2-6.0 Å) (32, 33). Similarly, K1544 is directed away from the IGF2 binding site in domain 11^{AB5}, but rotates to point towards it in domain $11^{AB5\text{-}Q1569R\ P1597H\ S1602H}$. Y1542 (AB loop) and F1567 (CD loop) showed minor conformational changes that open a deeper pocket at the F19 binding site than observed in the IGF2: domain11^{AB3} complex. Other residues in these loops adopted comparable conformations.

The FG loop of domain $11^{AB5\text{-}Q1569R\ P1597H\ S1602H}$ is well-defined and the side-chains of both mutated histidine residues were visible, revealing a significant rearrangement of this portion of the loop when compared to domain11^{AB5} (Fig. 4B). Whereas P1597 and S1602 do not interact in either the free or bound form of domain 11^{AB5}, the histidine residues associate closely in domain $11^{AB5\text{-}Q1569R\ P1597H\ S1602H}$ to produce a positively charged patch. Lastly, the HI loop backbone and side-chains superimpose well between domain 11^{AB5} and domain $11^{AB5\text{-}Q1569R\ P1597H\ S1602H}$. In terms of the complex with $IGF2^{1\text{-}67}$, we speculate that packing of W1546 into the hydrophobic core as seen with domain 11^{AB5} would require breaking the cation-π interaction with R1569. This would allow W1546 to interact with F19 and R1569 to form new electrostatic interactions with D15 and D23 of IGF2 (Fig. 4C). Secondly, the positively charged patch formed from H1597 and H1602 on the FG loop may interact with the N-terminus of $IGF2^{1\text{-}67}$ (e.g. E6) (Fig. 4D). S1600 and K1601, however, are also re-oriented favourably for interaction with IGF2 D52, a further key residue for binding on IGF2 (16). The position of I1572 and Y1606 does not occlude binding of T16 based on the domain 11^{AB3}:$IGF2^{1\text{-}67}$ complex and may contribute to the retained specificity for $IGF2^{1\text{-}67}$ over IGF1. These observations will require confirmation in future X-ray crystallography studies of the $IGF2^{1\text{-}67}$: domain $11^{AB5\text{-}Q1569R\ P1597H\ S1602H}$ complex (Fig. 4D).

Development of specific and soluble IGF2 antagonists based on domain 11^{AB5} mutants

As domain 11^{AB5}, and the combined variant domain $11^{AB5\text{-}Q1569R\ P1597H\ S1602H}$, had affinities that exceeded that of endogenous M6P/IGF2R, IGF1R, and IR-A, we evaluated whether these domains could be translated into the basis of improved IGF2 antagonists or ligand traps that function *in vivo* (14). Human IgG1 Fc-fusion domain 11^{AB5} and human the optimised IgG2 Fc-domain $11^{AB5\text{-}Q1569R\ P1597H\ S1602H}$ (from this point abbreviated to domain $11^{AB5\text{-}RHH}$) were bulk produced as initial optimised IGF2 antagonists, along with an IgG1 Fc-domain 11^{I1572A} control with mutation of the CD loop known to impair IGF2 binding (SI Appendix-Supplementary Fig. 7). SPR binding kinetics revealed an affinity and specificity for $IGF2^{1\text{-}67}$ close to that of the single soluble domain 11^{AB5} when IgG1 Fc-domain 11^{AB5} was immobilised using an anti-human Fc antibody (K_D = 3.89 nM, SI Appendix-Supplementary Table 9, Fig. 5A). $IGF2^{1\text{-}67}$ binding by IgG2 Fc-fusion domain $11^{AB5\text{-}RHH}$ also showed a significantly reduced k_{off} (0.0014 s^{-1}) and slower k_{on} to yield a similar overall K_D to the respective single domain (SI Appendix-Supplementary Table 9, Fig. 5A).

IGF2 in serum, and commonly in tumours, occurs as a range of pro-IGF2 isoforms attributed to incomplete processing of the E-domain by pro-protein convertases (34). Binding kinetics of recombinant pro-IGF2 isoforms (mature $IGF2^{1\text{-}67}$ [7.5 kDa], pro-$IGF2^{1\text{-}104}$ [11 kDa] and pro-$IGF2^{1\text{-}156}$ [17 kDa]) to Fc-domain 11^{AB5} and Fc-domain $11^{AB5\text{-}RHH}$, including specificity of binding relative to IGF1 were determined. We observed higher relative affinity of pro-$IGF2^{1\text{-}104}$ to Fc-domain $11^{AB5\text{-}RHH}$ than Fc-domain 11^{AB5}, and surprisingly binding of Fc-domain 11^{I1572A} to both pro-$IGF2^{1\text{-}156}$ and to a lesser extent pro-$IGF2^{1\text{-}104}$ (Fig. 5A, SI Appendix-Supplementary Table 9). Binding of domain 11^{AB5} and domain $11^{AB5\text{-}RHH}$ was independent of the Fc-fragment, as the same affinity to pro-IGF2 isoforms was obtained when immobilised human Fc alone was used in the control flow cell. No binding to IGF1 was observed (Fig. 5A). Solution pull-downs confirmed the specific capture of recombinant $IGF2^{1\text{-}67}$, pro-$IGF2^{1\text{-}104}$ and pro-$IGF2^{1\text{-}156}$ (note multiple forms) by Fc-domain 11^{AB5}, and binding of Fc-domain 11^{I1572A} to pro-$IGF2^{1\text{-}156}$ (Fig. 5B). In order to further test the functional activity of IGF binding, culture supernatants were obtained from human hepatocellular carcinoma cell lines (HuH7, Hep3B) and a Tet-inducible NIH-3T3 line that expressed big-IGF2 isoforms and $IGF2^{R104A}$, a mutation that impairs processing of $IGF2^{1\text{-}156}$ (34), respectively (Fig. 5C). Solution pull-downs from cell line supernatants confirmed Fc-domain 11^{AB5} binding to larger pro-IGF2 glycosylated isoforms, and domain 11^{I1572A} particularly to $IGF2^{1\text{-}156}$, indicating the potential for additional binding interactions of the E-domain of IGF2 to domain 11 independent of the main IGF2 binding surface.

Trapping IGF2 with Fc-domain $11^{AB5\text{-}RHH}$ (IGF2-TRAP) inhibits IGF2 dependent signalling *in vivo*

To test whether the domain 11^{AB5} backbone could function as an IGF2 antagonist or ligand trap *in vivo*, we tested the acute metabolic signalling effects of exogenous $IGF2^{1\text{-}67}$ induced hypoglycaemia in mice. Co-injection of IGF2 and Fc-domain 11^{AB5} or Fc-domain $11^{AB5\text{-}RHH}$ in either a 1 : 1, or 1 : 0.5 molar ratio, respectively, resulted in abrogation of the $IGF2^{1\text{-}67}$ induced hypoglycaemia, an effect not detected with the Fc-domain 11^{I1572A} non-IGF2 binding control (Fig. 6A, SI Appendix-Supplementary Fig. 8). When the IGF2 to Fc-domain 11 ratio was reduced, the advantages of a higher affinity mutant were evident. A molar ratio of 1 : 0.23 significantly counteracted IGF2-induced hypoglycaemia in the case of domain $11^{AB5\text{-}RHH}$, but not with Fc-domain 11^{AB5}. Moreover, pre-loading or post-loading of Fc-domain 11^{AB5} also abrogated $IGF2^{1\text{-}67}$ induced hypoglycaemia (SI Appendix-Supplementary Fig. 9).

Fc-domain 11^{AB5} inhibited cell growth of Hep3B and Huh7 cells in culture, but growth of these cell lines as xenografts failed in CD-1 nude mice *in vivo* (SI Appendix-Supplementary Fig. 10 and 11). An alternative autocrine-$IGF2^{1\text{-}67}$ tumour model was developed using the Ewing sarcoma cell line (SKNMC) selected for IGF2 dependency (35). Growth of retroviral transformed SKNMC cells expressing constitutive $IGF2^{1\text{-}67}$ and a luciferase reporter was inhibited by Fc-domain $11^{AB5\text{-}RHH}$ in culture, and inhibited IGF2 downstream signaling (SI Appendix-Supplementary Fig. 12). SKNMC-$IGF2^{1\text{-}67}$ xenografts are predicted to be dependent on autocrine-$IGF2^{1\text{-}67}$ supply, as adult mice do not express IGF2. IgG2 Fc-domain $11^{AB5\text{-}RHH}$ purified from CHO cells was infused using osmotic mini-pumps prior to SKNMC-$IGF2^{1\text{-}67}$ cell line injection. Reduced xenograft growth rates in mice were observed with Fc-domain $11^{AB5\text{-}RHH}$ infusion (Fig. 6B and SI Appendix-Supplementary Fig. 13A), with induced regions of tumour cell death (SI Appendix-Supplementary Fig. 13B) that resulted in reduced bioluminescence signal to volume ratio (SI Appendix-Supplementary Fig. 13C, D). Circulating $IGF2^{1\text{-}67}$ was readily detected in serum following 28 days of SKNMC-$IGF2^{67}$ tumour growth, but was less evident in treated Fc-domain $11^{AB5\text{-}RHH}$ mice, and without alteration in either total serum IGF1, GH or IGFBP levels (Fig. 6C).

The genomic complexity of cancer reduces any supposition that a drug targeting IGF2 alone in established tumours will lead to durable consequences, yet there may be specific circumstances where dual targeting with other agents may unmask a synergistic or conditional dependency on IGF2. A number of candidate additive and synergistic lethality interactions were observed in SKNMC-IGF2[104] treated with Fc-domain $11^{AB5\text{-}RHH}$ and an extended oncology compound library (222 agents plus 77 additional agents from the DTP Approved Oncology Drug Set contains oncology specific compounds that have been tested in man, http://www.tdi.ox.ac.uk/home). The high throughput screen compared the effects of the oncology drugs in the presence and absence of a fixed concentration of Fc-domain $11^{AB5\text{-}RHH}$ (IGF2-TRAP) that resulted in partial reduction in cell viability (10-20%) at the assay end point of 72hrs (SI Appendix-Supplementary Fig. 14A, B). Evaluation of the metrics of raw data (Pearson's r scores) indicated good replicate correlation, with average r ≥0.92 for all plates with an average inhibitory effect of the IGF2-TRAP alone of ∼10%. Drugs showing $p<0.05$ in the rank product method applied to their interaction score were selected. A number of hits were identified, including two independent PI3-kinase inhibitors, PF-04691502 ($p=0.011$) and Pictilisib ($p=0.017$) (Fig. 6D). These were further validated by comparison of the dose response curve both in the presence and absence of IGF2-TRAP (Fc-domain $11^{AB5\text{-}RHH}$) (Fig. 6E). The IC_{50} for PF-04691502 shifted from 0.42 μM to 92 nM in the presence of the IGF2-TRAP, and from 1.1 μM to 0.14 μM in the case of pictilisib, both of which are significant based on an F-test. Such ten-fold differences following combination with IGF2-TRAP suggest that antagonising IGF2 signalling can reveal synergistic activity that is likely to be clinically advantageous, at least in reducing the dosage of these PI3 kinase inhibitors.

Changes in the molecular distribution of circulating IGF2 upon IGF2-TRAP treatment were analysed by neutral size-fractionation of human serum (Fig. 6F). In adult human serum, western blots of the fractions reveal that IGF2 is mainly the mature lower molecular weight isoform, and is predominantly present in a ternary complex with IGFBP3 and the acid labile subunit (ALS) (~150 kDa), and to a lesser extent in a binary complex with IGFBP3 (~40 kDa), with free IGF2 below the limit of detection. In this sample, IGF2-TRAP depleted IGF2 in the binary complex, and also partially decreased the amount IGF2 in the ternary complex (lower left panel, Fig. 6F). Pull-downs confirmed that the depleted IGF2 specifically bound IGF2-TRAP. In a second experiment, serum from a patient with non-islet-cell-tumour-associated hypoglycaemia (NICTH) secondary to a GIST tumour was utilised. Western blotting of the fractions showed that IGF2 was present with additional larger molecular weight isoforms, predominantly in a binary complex. IGF2-TRAP depleted both lower and higher molecular weight isoforms of IGF2, and pull-downs confirmed specific binding to both isoforms (lower right panel, Fig. 6F). These data suggest that IGF2-TRAP can target larger molecular weight pro-IGF2 isoforms associated with human cancer and NICTH (IGF2 syndrome).

Discussion

Functional evolution of affinity without negative consequences on specificity

When comparing the binding kinetics in the context of high resolution NMR structures of the domain 11^{WT}, domain 11^{AB3} and the new structure domain 11^{AB5}, structural perturbations within the binding site are mainly localised to the AB loop (2). Although the G1546W mutation in domain 11^{AB5} introduces a bulky side-chain into the AB loop and the binding site, residues in the neighbouring CD and FG loops appear only minimally perturbed in this structure. The exciting structural and functional observation is, however, that affinity for IGF2 can be increased by further ligand interacting surface and rim (non-AB loop) mutations that confer re-configuration of the binding surface complementarity, rather than acting indirectly through non-ligand interacting surface sites that are remote (40 Å) from the interface (36). The predominant polar and electrostatic mutations are consistent with the potential for both ligand-loop stabilisation of existing hotspot interactions through second sphere residues, as the specific effects are mainly selected on k_{off} and the structural complementarity surrounding F19, L53 and T16 of IGF2. Importantly, this supports the maintained selectivity through T16 of IGF2, so retaining specificity over IGF1. The thermodynamic pay-offs between entropy (predominantly solvent exclusion) and enthalpy (charge attraction) of the interaction between IGF2 and domain 11 mutants showed some consistent differences, including the effects of CD and FG loop mutations. Mutation of the AB loop in the domain 11^{AB3} would be consistent with rigidification and improved solvent exclusion, but this effect was only unmasked in domain 11^{AB5} with respect to thermodynamic terms after the introduction of additional mutations in the CD and FG loops (domain $11^{AB5\text{-}RHH}$). Thus, the overall improvement in Gibbs free energy of the interaction (ΔG^o) appeared to require combinations of mutations that reduced the entropic barrier (improved solvent exclusion), offsetting the negative cooperativity of the AB loop mutations. Ultimately, the combinatorial gain of function mutations in the AB, CD and FG loops in domain$11^{AB5\text{-}RHH}$ yield a sub-nanomolar IGF2 affinity that is exceptional for non-Ig scaffolds (37). This scaffold includes the unexpected abstraction of AB loop W1546 via R1569 from the IGF2 binding site that significantly modifies the IGF2 interaction, and the E6 IGF2 interaction with the FG intra-loop histidine-histidine bridge. These molecular mechanisms could never have been predicted even with prior structural knowledge, but were ultimately only discovered through integration of mutagenesis, selection and structure based re-engineering. Whether changing the permissive background of domain 11 binding site amino acids would have ultimately altered the specific selection of high affinity mutations is not known, but likely, and has implications for the modification of other ligand-receptor interactions (38).

Similar to mutations introduced into other protein-protein interaction sites, our data suggest that single point mutations in loops CD, HI and FG can result in relatively small (<2 fold) incremental gains in affinity in a domain isolated from the natural context of the receptor. It is also evident that, although such small gains are possible, they have not been selected in the current mammalian evolutionary context, indicating they have had little consequence on IGF2 and M6P/IGF2R function. The larger incremental gains in affinity of 8-10 fold observed during the evolution of early mammals (ancestors of monotremes) do, however, suggest that loop mutations have had a very significant impact on gain in fitness in the past (2). Combinatorial mutations in separate loops that result in much higher affinities would have a very low stochastic likelihood of occurring simultaneously *in vivo*. The five mutations of domain 11^{AB5} may therefore have been sufficient to generate a significant incremental gain function, and selected for potential improvement in fitness over time, if they could be introduced.

A functional high affinity and specificity IGF2 soluble antagonist

Activation of the insulin-like growth factor signaling pathway in cancer predominantly relates to gain of function of IGF2 ligand (expression and bio-availability), and the subsequent activation of both IGF1 and IR-A signaling receptors (39). Multiple mechanisms result in increased IGF2 expression and ligand supply, from genomic amplification to proteolysis of IGF binding proteins (40-46). Despite intensive efforts, sole targeting of IGF1R downstream of IGF2 appears limited in success in clinical trials (47). One problem appears the under-estimated effects of

IGF1(R) blockade on activation of a central nervous system negative feedback loop resulting in increased systemic GH driving total IGF1 production from the liver, GH mediated insulin resistance and increased insulin supply, resulting in antagonism of the IGF1R inhibitor and systemic metabolic toxicity (20). Similar compensatory feedback responses are also observed with anti-IGF ligand antibodies when administered at therapeutic levels, probably because of the lack of selectivity and subsequent sequestration of free IGF1 (48, 49). Anti-ligand antibodies appear to bind IGF2 with cross-reaction with IGF1, e.g. Dx-2647 (IGF2 49 pmol.L^{-1} vs IGF1 9.6 nmol.L^{-1}), Medi-573 (IGF2 2 pmol.L^{-1} vs IGF1 294 pmol.L^{-1}), m660 chimeric antibody (IGF2 700 pmol.L^{-1} vs IGF1 8 nmol.L^{-1}) (14, 50-52). For ligand trap molecules to be functional, it is well established a >10-fold high affinity than the affinity for the signaling receptor appears necessary in order to achieve maximal antagonistic activity (53). A specific IGF2 antagonist such as Fc-domain $11^{AB5\text{-}RHH}$ (IGF2-TRAP) may have these significant advantages, because its higher specific affinity means it can inhibit IGF2 activation of both IGF1R and IR-A signaling at the level of the cell, but may also have the potential advantage of not perturbing the endocrine hypothalamic IGF1-GH-insulin feedback that occurs when targeting IGF1. The latter, however, would require evaluation in human trials, as humans have an adult liver *IGF2* promoter, resulting in gene expression and circulating ligand.

Synergistic effects with PI3 kinase inhibitors shown here, are also in keeping with reports of the intrinsic resistance conferred by the IGF pathway following PI3 kinase inhibition (ZSTK474) in gastric carcinoma cell lines, suggesting conditional effects acting directly on the intra-cellular signaling pathway may unmask such IGF1R/IR-A and IGF2 ligand dependency (54). Similar intra-cellular mechanisms of feedback may extend to regulation of ligand bio-availability, as EGFR pathway resistance may be conferred by reduced IGFBP levels, and resistance to IGF pathway inhibition may be improved by downstream MEK inhibition (55, 56). Our data supports specific IGF2 targeting might also address unmet needs for IGF2 dependent clinical phenotypes, such as the mosaic over-growth Beckwith-Weidemann syndrome (GCID: GC11M002113), and as stated, the systemic excess of larger molecular weight or big-IGF2 ($IGF2^{87}$, $IGF2^{104}$) in NICTH (IGF2 syndrome) (57, 58). Our functional data suggest that IGF2 signaling, with and without other targeted cell signaling agents, can be inhibited by the high affinity and specificity generated by combinatorial mutation in Fc-domain $11^{AB5\text{-}RHH}$ (IGF2-TRAP) *in vivo*.

In summary, we have rapidly evolved the domain 11 IGF2 binding site and identified novel combinatorial molecular and structural mechanisms of enhanced affinity that are confined to the surface of the binding interface. Specifically, these include inter-domain (AB-CD) and intra-domain (FG-FG) amino acid side chain interactions that alter the complimentary interactions with IGF2, but that retain IGF2 T16 interactions that confer specificity over IGF1. The 100-fold improvement in IGF2: domain 11 affinity and our functional assays of ligand trap activity, including binding to pathological pro-IGF2 isoforms, accelerate the impetus to apply the advantages of these novel domains as IGF2 ligand traps for cancer therapy.

Materials and Methods

(Further details are reported in SI Appendix-Supplementary Material and Methods).

Yeast surface display of domain 11 and IGF2 binding screens.

The yeast surface display of a N-terminal Flag tagged fusion of domain 11 and the C-terminal domain of α-agglutinin was developed in *P. pastoris* for loop-specific mutant library selection using binding with biotinylated IGF2 and flow sorting (SI Appendix-Supplementary Table 1) as described (59). Selected mutant proteins were purified by Ni-NTA column prior to validation by SPR on a T200 Biacore. Thermodynamic data was performed across the temperature range of 10 °C to 30 °C, van't Hoff and Erying plots generated and ΔH, $-T\Delta S$ and ΔG values determined as previously described (22).

Structure-informed loop-specific site directed mutagenesis.

The domain 11^{AB3}:IGF2 structure (PDB: 2L29) was used to identify residues in the binding loops that could interact with IGF2. These were manually mutated in Pymol v1.5 to determine whether or not they might stabilise the complex (http://www.pymol.org/). In addition the Robetta server (http://robetta.bakerlab.org/alascansubmit.jsp) was used to interface interface alanine scanning mutagenesis on the AB3:IGF2 complex.

Expression and purification of domain 11^{WT}, domain 11^{AB5} and IGF2-domain 11^{AB5} complexes for NMR and ^{15}N relaxation studies of domain11^{WT} and domain 11^{AB5}.

His_6-tagged domain 11^{WT} and domain 11^{AB5} were isotopically ^{15}N or dual ^{15}N, ^{13}C-labeled and purified from *E. coli* (2). Structural calculations of the domain 11AB5 both free and complexed with IGF2 were as described previously (2). NOE, ^{15}N-T_1 and ^{15}N-T_2 NMR relaxation data were acquired at two fields (600 and 900 MHz) at 25 °C. T_1 and T_2 experiments were run in duplicate (SI Appendix-Supplementary Fig. 2.

Expression, purification and crystal structure of domain $11^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$

Domain $11^{AB5\text{-}\ Q1569R\ P1597H\ S1602H}$ protein (160 μM) screened for crystals and data collection carried out using PILATUS detector at station I04 of the Diamond synchrotron (SI (Appendix-Supplementary Table 8).

Fc-fusion protein expression and purification.

IgG1 Fc-domain 11^{AB5} fusion protein was produced by transient transfection of HEK293T cells and IgG2 Fc-domain $11^{AB5\text{-}RHH}$ in Chinese Hamster Ovary GS knockout cells and purified by protein A affinity (SI Appendix-Supplementary Fig. 7).

IGF2 induced hypoglycaemia.

All animal experiments were approved by UK Home Office (PPL 30/2695). Blood glucose levels normalised for 30 min after induction of anaesthesia prior to IGF2 injection either alone or premixed with different molar ratios Fc-domain 11^{AB5} or Fc-domain 11^{I1572A}.

IGF2 dependent human cancer cell xenografts.

SKNMC-$IGF2^{67}$ cells mixed with matrigel were injected into CD-1® (Crl:NU-1 *Foxn1^{nu}*), 24 hours after ALZET® mini-pump placement delivering 40 mg kg^{-1} per week either IGF2-TRAP or PBS control. Tumour growth was monitored with both caliper measurements and bioluminescence.

IGF2-TRAP synergistic lethality screens.

SKNMC-$IGF2^{104}$-IRES-luc cells were exposed to a 307 oncology drug library that have been tested in man (TDI expanded Oncology library) using10 μM, 1 μM, 100 nM and 10 nM drug, in the presence and absence of 250 nM IGF2-TRAP.

Neutral size-fractionation of human serum and IGF2 depletion with IGF2-TRAP.

Human serum was size-fractionated by gel filtration and western blot for the presence of ALS, IGFBP3 and IGF2, and pull-downs performed using protein G magnetic beads (see SI Appendix-Supplementary Material and Methods).

Acknowledgements.

We thank Cancer Research UK (ABH C429), EPSRC, Algerian Government (DR), NIHR Biomedical Research Centre and Unit, Oxford, Clarendon Fund and Agilent Technologies (formerly Varian) for support, and the Wellcome Trust for the use of 900 MHz NMR spectrometer at the HWB-NMR, Birmingham, UK. We thank Emily Richards, Leanne Minnall and Sara Whittaker for NMR data, Hans-Jürgen Hoppe, Sermet Can, Claudia Bühnemann, Nigel Rust, Martin Glennie, Anton van der Merwe and Neil Barclay for discussion. The research described in this manuscript is the subject of a worldwide patent application WO2007020402A, granted in Europe as EP1945663B and pending in the US as US20090304651A. The ensembles of NMR structures and associated NMR chemical shifts have been deposited with the protein database and BioMagResBank: Free domain 11^{AB5} (2M6T and rcsb103281), IGF2 bound domain 11^{AB5} (2M68 and rcsb103260). The crystal structure of domain $11^{AB5\text{-}RHH}$ has been deposited with PDB code 5IEI. **Author Contributions** SF, DR, OZ and ABH devised the yeast directed evolution strategy; SF and DR identified domain11^{AB5} and performed yeast surface display mutants. MS, CW, and MPC determined the NMR structure and dynamics of domain11^{AB5}; RN, CW, MS and MPC expressed structural directed mutants; SF performed all SPR; LG and SF generated Fc fusion proteins; SF and RN generated, and SF characterised combination mutants; RN, CW and MPC determined structures of domain$11^{AB5\text{-}RHH}$. SF, JH and MiS performed hypoglycemia experiments; SF, RM, SS, DE performed cellular screens, tumour and protein assays; DE, SY and FB assisted in synergistic screens, SF and ABH xenograft experiments; SF, MS, CW, MPC and ABH wrote the manuscript. **Conflict of Interest** The authors declare no conflict of interest.

1. Harms MJ & Thornton JW (2013) Evolutionary biochemistry: revealing the historical and

physical causes of protein properties. *Nat Rev Genet* 14(8):559-571.
2. Williams C, *et al.* (2012) An exon splice enhancer primes IGF2:IGF2R binding site structure and function evolution. *Science* 338(6111):1209-1213.
3. Ghosh P, Dahms NM, & Kornfeld S (2003) Mannose 6-phosphate receptors: new twists in the tale. *Nat Rev Mol Cell Biol* 4(3):202-212.
4. Linnell J, Groeger G, & Hassan AB (2001) Real time kinetics of insulin-like growth factor II (IGF-II) interaction with the IGF-II/mannose 6-phosphate receptor: the effects of domain 13 and pH. *J Biol Chem* 276(26):23986-23991.
5. Morgan DO, *et al.* (1987) Insulin-like growth factor II receptor as a multifunctional binding protein. *Nature* 329(6137):301-307.
6. Wilkins JF & Haig D (2003) What good is genomic imprinting: the function of parent-specific gene expression. *Nat Rev Genet* 4(5):359-368.
7. Hughes J, Frago S, Buhnemann C, Carter EJ, & Hassan AB (2013) Maternal transmission of a humanised Igf2r allele results in an Igf2 dependent hypomorphic and non-viable growth phenotype. *PLoS One* 8(2):e57270.
8. Lau MM, *et al.* (1994) Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes Dev* 8(24):2953-2963.
9. Wang ZQ, Fung MR, Barlow DP, & Wagner EF (1994) Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mpr gene. *Nature* 372(6505):464-467.
10. Wylie AA, *et al.* (2003) Tissue-specific inactivation of murine M6P/IGF2R. *Am J Pathol* 162(1):321-328.
11. De Souza AT, Hankins GR, Washington MK, Orton TC, & Jirtle RL (1995) M6P/IGF2R gene is mutated in human hepatocellular carcinomas with loss of heterozygosity. *Nat Genet* 11(4):447-449.
12. Hankins GR, *et al.* (1996) M6P/IGF2 receptor: a candidate breast tumor suppressor gene. *Oncogene* 12(9):2003-2009.
13. Harper J, *et al.* (2006) Soluble IGF2 receptor rescues Apc(Min/+) intestinal adenoma progression induced by Igf2 loss of imprinting. *Cancer Res* 66(4):1940-1948.
14. Prince SN, Foulstone EJ, Zaccheo OJ, Williams C, & Hassan AB (2007) Functional evaluation of novel soluble insulin-like growth factor (IGF)-II-specific ligand traps based on modified domain 11 of the human IGF2 receptor. *Mol Cancer Ther* 6(2):607-617.
15. Surinya KH, *et al.* (2008) An investigation of the ligand binding properties and negative cooperativity of soluble insulin-like growth factor receptors. *J Biol Chem* 283(9):5355-5363.
16. Delaine C, *et al.* (2007) A novel binding site for the human insulin-like growth factor-II (IGF-II)/mannose 6-phosphate receptor on IGF-II. *J Biol Chem* 282(26):18886-18894.
17. Bach LA, Headey SJ, & Norton RS (2005) IGF-binding proteins--the pieces are falling into place. *Trends in endocrinology and metabolism: TEM* 16(5):228-234.
18. Greenall SA, *et al.* (2013) Biochemical characterization of individual human glycosylated pro-insulin-like growth factor (IGF)-II and big-IGF-II isoforms associated with cancer. *J Biol Chem* 288(1):59-68.
19. Zaina S & Squire S (1998) The soluble type 2 insulin-like growth factor (IGF-II) receptor reduces organ size by IGF-II-mediated and IGF-II-independent mechanisms. *J Biol Chem* 273(44):28610-28616.
20. Yee D (2012) Insulin-like growth factor receptor inhibitors: baby or the bathwater? *J Natl Cancer Inst* 104(13):975-981.
21. Brown J, *et al.* (2008) Structure and functional analysis of the IGF-II/IGF2R interaction. *EMBO J* 27(1):265-276.
22. Zaccheo OJ, *et al.* (2006) Kinetics of insulin-like growth factor II (IGF-II) interaction with domain 11 of the human IGF-II/mannose 6-phosphate receptor: function of CD and AB loop solvent-exposed residues. *J Mol Biol* 359(2):403-421.
23. Canfield WM & Kornfeld S (1989) The chicken liver cation-independent mannose 6-phosphate receptor lacks the high affinity binding site for insulin-like growth factor II. *J Biol Chem* 264(13):7100-7103.
24. Clairmont KB & Czech MP (1989) Chicken and Xenopus mannose 6-phosphate receptors fail to bind insulin-like growth factor II. *J Biol Chem* 264(28):16390-16392.
25. Yandell CA, Dunbar AJ, Wheldrake JF, & Upton Z (1999) The kangaroo cation-independent mannose 6-phosphate receptor binds insulin-like growth factor II with low affinity. *J Biol Chem* 274(38):27076-27082.
26. Bloom JD & Arnold FH (2009) In the light of directed evolution: pathways of adaptive protein evolution. *Proc Natl Acad Sci U S A* 106 Suppl 1:9995-10000.
27. Levin KB, *et al.* (2009) Following evolutionary paths to protein-protein interactions with high affinity and selectivity. *Nat Struct Mol Biol* 16(10):1049-1055.
28. Yuen CM & Liu DR (2007) Dissecting protein structure and function using directed evolution. *Nat Methods* 4(12):995-997.
29. Mergler M, Wolf K, & Zimmermann M (2004) Development of a bisphenol A-adsorbing yeast by surface display of the Kluyveromyces yellow enzyme on Pichia pastoris. *Appl Microbiol Biotechnol* 63(4):418-421.
30. Williams C, *et al.* (2007) Structural insights into the interaction of insulin-like growth factor 2 with IGF2R domain 11. *Structure* 15(9):1065-1078.
31. Tinoco I, Sauer K, & Wang JC (1978) *Physical chemistry : prinicples and applications in biological sciences* (Prentice-Hall, Englewood Cliffs, N.J.) pp xv, 624 p.
32. Dougherty DA (2007) Cation-pi interactions involving aromatic amino acids. *The Journal of nutrition* 137(6 Suppl 1):1504S-1508S; discussion 1516S-1517S.
33. Gallivan JP & Dougherty DA (1999) Cation-pi interactions in structural biology. *Proc Natl Acad Sci U S A* 96(17):9459-9464.
34. Duguay SJ, *et al.* (1998) Post-translational processing of the insulin-like growth factor-2 precursor. Analysis of O-glycosylation and endoproteolysis. *J Biol Chem* 273(29):18443-18451.
35. Buhnemann C, *et al.* (2014) Quantification of the heterogeneity of prognostic cellular biomarkers in ewing sarcoma using automated image and random survival forest analysis. *PLoS One* 9(9):e107105.
36. Kastritis PL, Rodrigues JP, Folkers GE, Boelens R, & Bonvin AM (2014) Proteins feel more than they see: fine-tuning of binding affinity by properties of the non-interacting surface. *J Mol Biol* 426(14):2632-2652.
37. Skrlec K, Strukelj B, & Berlec A (2015) Non-immunoglobulin scaffolds: a focus on their targets. *Trends Biotechnol*.
38. Harms MJ & Thornton JW (2014) Historical contingency and its biophysical basis in glucocorticoid receptor evolution. *Nature* 512(7513):203-207.
39. Brouwer-Visser J & Huang GS (2015) IGF2 signaling and regulation in cancer. *Cytokine Growth Factor Rev*.
40. Chess A (2012) Mechanisms and consequences of widespread random monoallelic expression. *Nat Rev Genet* 13(6):421-428.
41. Gaudet F, *et al.* (2003) Induction of tumors in mice by genomic hypomethylation. *Science* 300(5618):489-492.
42. Lui JC & Baron J (2013) Evidence that Igf2 down-regulation in postnatal tissues and up-regulation in malignancies is driven by transcription factor E2f3. *Proc Natl Acad Sci U S A* 110(15):6181-6186.
43. Miyamoto S, *et al.* (2007) Matrix metalloproteinase-7 triggers the matricrine action of insulin-like growth factor-II via proteinase activity on insulin-like growth factor binding protein 2 in the extracellular matrix. *Cancer Sci* 98(5):685-691.
44. Seshagiri S, *et al.* (2012) Recurrent R-spondin fusions in colon cancer. *Nature* 488(7413):660-664.
45. Sun Y, *et al.* (2006) IGF2 is critical for tumorigenesis by synovial sarcoma oncoprotein SYT-SSX1. *Oncogene* 25(7):1042-1052.
46. Vidal SJ, *et al.* (2015) A targetable GATA2-IGF2 axis confers aggressiveness in lethal prostate cancer. *Cancer Cell* 27(2):223-239.
47. Yee D (2015) A tale of two receptors: insulin and insulin-like growth factor signaling in cancer. *Clin Cancer Res* 21(4):667-669.
48. Haluska P, *et al.* (2014) Phase I dose-escalation study of MEDI-573, a bispecific, antiligand monoclonal antibody against IGFI and IGFII, in patients with advanced solid tumors. *Clinical cancer research : an official journal of the American Association for Cancer Research* 20(18):4747-4757.
49. Mireuta M, Birman E, Barmash M, & Pollak M (2014) Quantification of binding of IGF-1 to BI 836845, a candidate therapeutic antibody against IGF-1 and IGF-2, and effects of this antibody on IGF-1:IGFBP-3 complexes in vitro and in male C57BL/6 mice. *Endocrinology* 155(3):703-715.
50. Gao J, *et al.* (2011) Dual IGF-I/II-neutralizing antibody MEDI-573 potently inhibits IGF signaling and tumor growth. *Cancer Res* 71(3):1029-1040.
51. Chen W, Feng Y, Zhao Q, Zhu Z, & Dimitrov DS (2012) Human monoclonal antibodies targeting nonoverlapping epitopes on insulin-like growth factor II as a novel type of candidate cancer therapeutics. *Mol Cancer Ther* 11(7):1400-1410.
52. Dransfield DT, *et al.* (2010) A human monoclonal antibody against insulin-like growth factor-II blocks the growth of human hepatocellular carcinoma cell lines in vitro and in vivo. *Mol Cancer Ther* 9(6):1809-1819.
53. Forsten KE & Lauffenburger DA (1992) Interrupting autocrine ligand-receptor binding: comparison between receptor blockers and ligand decoys. *Biophysical journal* 63(3):857-861.
54. Isoyama S, *et al.* (2015) Basal expression of insulin-like growth factor 1 receptor determines intrinsic resistance of cancer cells to a phosphatidylinositol 3-kinase inhibitor ZSTK474. *Cancer Sci* 106(2):171-178.
55. Flanigan SA, *et al.* (2013) Overcoming IGF1R/IR resistance through inhibition of MEK signaling in colorectal cancer models. *Clinical cancer research : an official journal of the American Association for Cancer Research* 19(22):6219-6229.
56. Guix M, *et al.* (2008) Acquired resistance to EGFR tyrosine kinase inhibitors in cancer cells is mediated by loss of IGF-binding proteins. *J Clin Invest* 118(7):2609-2619.
57. Fukuda I, *et al.* (2006) Clinical features of insulin-like growth factor-II producing non-islet-cell tumor hypoglycemia. *Growth Horm IGF Res* 16(4):211-216.
58. Rikhof B, *et al.* (2009) Insulin-like growth factors and insulin-like growth factor-binding proteins in relation to disease status and incidence of hypoglycaemia in patients with a gastrointestinal stromal tumour. *Ann Oncol* 20(9):1582-1588.
59. Wu S & Letchworth GJ (2004) High efficiency transformation by electroporation of Pichia pastoris pretreated with lithium acetate and dithiothreitol. *Biotechniques* 36(1):152-154.

Supporting Information- Appendix

Functional evolution of IGF2:IGF2R domain 11 binding generates novel structural interactions and a specific IGF2 antagonist

Susana Frago[1], Ryan D. Nicholls[2], Madeleine Strickland[2], Jennifer Hughes[1], Christopher Williams[2], Lee Garner[1], Mirvat Surakhy[1], Rory Mclean[1],Dellel Rezgui[1], Stuart N. Prince[1], Oliver J. Zaccheo[1], Daniel Ebner[3], Sabina Sanegre[1], Sheng Yu[4], Francesca Buffa[4], Matthew P. Crump[2¶] and A. Bassim Hassan[1¶].

Material and Methods

Yeast surface display of domain 11 and IGF2 binding screens.

The yeast surface display vector, pPICFlagDomain 11-Agα1, was generated by ligating IGF2R domain 11 (residues 1508-1654) into the *Eco*RI and *Avr*II restriction sites of pPIC9K preceded by the Flag peptide and followed by the C-terminal region of the *Saccharomyces cerevisiae AGα1* gene. The resulting protein is a N-terminal Flag tagged fusion of domain 11 and the C-terminal domain (residues 311-650) of α-agglutinin, which targets the fusion protein to the cell surface through the secretory pathway, where it is attached by a GPI anchor (1). The pLITMUS-Domain 11 vector was also generated by ligation of the Domain 11 insert into the *Eco*RI and *Avr*II restriction sites in the multipurpose cloning pLITMUS28i vector. Loop-specific mutant libraries were generated by either whole plasmid PCR (using Kappa HiFi HotStart DNA polymerase, Kappa Biosystems, or Phusion Polymerase, Finnzymes), using tail-to-tail mutagenic primers pairs (Supplementary Table 1) and the pPICFlagDomain11-Agα1 vector or the pLITMUS-Domain 11 vector as template for the AB loop and the HI loop libraries, respectively. The PCR product was treated with *Dpn*I, self-ligated and transformed into XL10-Gold® ultra-competent *E. coli* cells (Stratagene) to generate the *E. coli* libraries. The presence of random amino acid mutations in the desired positions was confirmed by sequencing. In the case of the HI loop library, a midiprep of the *E. coli* pLITMUS-Domain 11^{HI} library (containing 8.3×10^6 clones) was digested with *Eco*RI and *Avr*II and the excised HI loop-mutated domain 11 was then ligated into the pPICFlag-Agα1 and transformed into the ultra-competent *E. coli* cells to generate the *E. coli* pPICFlagDomain 11^{HI}-Agα library (containing 1.32 x 10^6 mutants). pPICFlagDomain 11^{AB}-Agα1 and pPICFlagDomain 11^{HI}-Agα1 were linearised within the HIS4 gene by *Sal*I digestion and electroporated into the histidine auxothrophic *P. pastoris* strain GS115 as described by (2). For selection, *P. pastoris* libraries were grown in histidine deficient minimal medium (MD=1.34% yeast nitrogen base (YNB), 4×10^{-5} % biotin, 2% dextrose).

For surface display expression, *P. pastoris* yeast libraries selected on MD for 72 hr, were kept at 4 ºC for a further 24 hr and passaged on MD. An excess of at least 10-fold of the library size of freshly grown cells was grown at 30ºC in BMGY (1% yeast extract, 2% peptone, 1.34% YNB, 4 x 10-5% biotin, 1% (v/v) glycerol, 100 mM potassium phosphate, pH 6.0) for 24-48 hr. Surface display of the mutated domain 11 was induced by transfer to BMMY (1% yeast extract, 2% peptone, 1.34% YNB, 4 × 10-5% biotin, 0.5% (v/v) methanol, 100 mM potassium phosphate, pH 6.0). For maintained induction, cultures were supplemented with

methanol to a concentration of 1% (v/v) every 24 hr for three days. Yeast cultures of control proteins (domain 11^{WT}, domain 11^{AB3} and domain 11^{I1572A}) were induced in the same manner.

Freshly induced yeast cells displaying domain 11^{WT}, domain 11^{AB3} and domain 11^{I1572A} were pelleted and washed with PBSA (PBS+0.5% (v/v) BSA). Aliquots of 5×10^5 cells were incubated in PBSA for 1 hr at room temperature with anti-Flag M2 antibody (Sigma) and increasing concentrations of biotinylated IGF2 ranging from 15 pM to 1.5 μM in a total volume such that a 10-fold molar excess of IGF2 over the domain 11 displayed on the surface was maintained (assuming a maximum of $5x10^4$ molecules per cell) (3). The cells were then pelleted, washed 3 times with cold PBSA and labeled with Alexa Fluor 488 goat anti-mouse IgG1 (Invitrogen) and Alexa Fluor® 647 streptavidin (Invitrogen) in a volume of 50 μL for 30 min at 4 ºC. Negative controls with secondary only and a positive control with anti-Flag+Alexa Fluor 488 only were also included. The cells were pelleted, washed 3 times with cold PBSA and re-suspended in 1 mL of PBSA and run in CyAn ADP Cytometer (Dako). The yeast population was gated on forward- and side-scatter channels to remove debris and the percentage of double positive cells was measured for each IGF2 concentration.

For the yeast loop libraries sorting, a 10-fold excess of the library population diversity was labeled in a similar manner, alongside the yeast cells displaying domain 11^{AB3}. The cells were pelleted, washed with PBSA and incubated first with the Anti-Flag M2 antibody (Sigma) for 1 hr at room temperature followed by 3x PBSA washes. The subsequent incubation with 5 nM IGF2, maintaining a 10-fold molar excess of IGF2 over domain 11, was also carried out for 1hr at room temperature. After 3x PBSA washes, the yeast cells were incubated with Alexa Fluor 488 goat anti-mouse IgG1 (Invitrogen) and Alexa Fluor® 647 streptavidin, Invitrogen) for 30 min-1 hr at 4 ºC. Negative and positive controls were similarly labeled and used, together with the yeast cells displaying domain 11^{AB3} (also labeled with 5 nM IGF2), to gate the library yeast cells that should display higher affinity binders on the surface (3).

FACS gated cells were grown for 48 hr in YPD (1% yeast extract, 2% peptone, 2% dextrose) containing 100U mL^{-1} penicillin, 100 μg mL^{-1} streptomycin, 100 μg mL^{-1} kanamycin and 0.5 mg mL^{-1} Geneticin. Genomic DNA was extracted using the YeaStar Genomic DNA kit (Zymoresearch). Mutated domain 11 was amplified by PCR from the genomic DNA using primers that incorporated *Eco*RI and *Avr*II restriction sites into the pPIC-HIS vector in frame with the α-factor secretion signal to generate pPIC-Domain 11 construct of the mutants (4). The vector was then linearised with *Sal*I and integrated into the HIS4 locus of the *P. pastoris* genome by electroporation (2). Yeast cells incorporating domain

11 into their genome were selected at 30 ºC on MD plates. Single yeast colonies were inoculated into 96 well plates containing 500 µl of BMGY and grown at 30 ºC for 24 hr (covered with a sterile breathable seal). Each plate contained a set of controls, including non transformed cells and cells secreting domain11^{WT}, domain 11^{AB3} and domain 11^{I1572A} mutants. The medium was replaced with BMMY and induction was sustained by supplementing with methanol to a final concentration of 1% (v/v) every 24 hr for two days.

Supernatants (100 µl) containing the secreted domain 11 mutants, were transferred to a 96 well plate, diluted with 100 µL of HBS-EP (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.005% (v/v) surfactant P20) and screened for IGF2 binding ability by SPR on a Biacore 3000. Biotinylated IGF2 (GroPep, Aus) was immobilized on the flow cells of the sensor chip SA by affinity capture to streptavidin. Kinetic experiments were performed at 25 ºC in HBS-EP at a flow-rate of 40 µL min^{-1} and consisted of a 2 min injection of diluted supernatant followed by a 2 min dissociation phase in HBS-EP running buffer, after which the binding surface was regenerated with a 2 min injection of 2M $MgCl_2$. The resulting sensorgrams were visually analysed using the standard BiaEvaluation software version 4.0.1 for Biacore 3000. Additionally, the dissociation phase was fitted to a first order exponential decay curve ($Y=Y_0 \cdot e^{-k_{off} \cdot x}$) from which the k_{off} could be calculated.

Protein expression and purification of M6P/ IGF2R domain 11 mutants for SPR studies.

For expression, yeast colonies selected on MD plates were grown at 30 ºC in 10 mL of BMGY for 24 hr and induced by being transferred to 50 mL BMMY. For maintained induction, cultures were supplemented with methanol to a final concentration of 1% (v/v) every 24 hr for two days. Supernatants were then subjected to SDS-PAGE and western blot using anti-His_6 mouse monoclonal antibody conjugated to peroxidase (Roche Diagnostics, USA). Domain 11 secreting cultures were scaled up to 500 mL BMMY and grown for further 72 hr and supplemented with 1% methanol every 24 hr. Supernatants were cleared by centrifugation (6000rpm for 20 min at 4 ºC in a Beckman-Coulter Avanti-J2 refrigerated centrifuge), diluted 5-fold in 20 mM sodium phosphate pH 8.0 buffer and His-purified using a Ni-NTA superflow column (Qiagen) on an Äkta FPLC system (GE Life Sciences, UK) washed with 50 mM imidazole and eluted with 250 mM imidazole. The fractions containing the protein were pooled and, when required, diluted in 20 mM MES pH 6.0 and purified using a Resource S column on an Äkta FPLC system (GE Life Sciences, UK). Purified proteins were dialyzed and concentrated using centrifugal filter units of 5 kDa cut-off (Millipore). Protein concentration was determined spectrophotometrically (Nanodrop) using

the theoretical extinction coefficient calculated by ProtParam and protein size and homogeneity was verified using 12% SDS-PAGE.

Structure-informed loop-specific site directed mutagenesis.

The domain 11^{AB3}:IGF2 structure (PDB: 2L29) was used to identify residues in the binding loops that could interact with IGF2. These were manually mutated in Pymol v1.5 to determine whether or not they might stabilise the complex (http://www.pymol.org/). In addition the Robetta server (http://robetta.bakerlab.org/alascansubmit.jsp) was used to perform interface alanine scanning mutagenesis on the Domain11^{AB3} complex to identify residues important for complex formation. A number of possible amino acids were introduced at each site to assess the residue's role in binding. M6P/IGF2R domain 11 sub-cloned from the pPICHis vector into pET26a (Novagen, Merck Chemicals Ltd, Nottingham, UK) was used as a template for site directed mutagenesis. The template was amplified in a thermocycler (PeqStar) using mutagenic primers and KOD DNA polymerase (Novagen) to introduce the mutations. Following amplification the parental DNA was digested with *DpnI* and transformed into *E. coli* 5α cells (NEB) for sequencing. The mutant proteins were expressed and refolded using existing protocols in *E. coli* BL21 (DE3) before the binding kinetics were analysed by SPR (5).

Expression and purification of domain 11^{WT}, domain 11^{AB5} and IGF2-domain 11^{AB5} complexes for NMR.

His_6-tagged domain 11^{WT} and domain 11^{AB5} (residues 1508-1654) cloned into pET26 (Novagen, Merck Chemicals Ltd, Nottingham, UK) were transformed into *E. coli* BL21(DE3) CodonPlus™ competent cells, purified as inclusion bodies and refolded according to well-established protocols (5). Proteins were isotopically labelled using 1 g L^{-1} ^{15}N-ammonium chloride and/or 2 g L^{-1} $^{13}C_6$-glucose (99%, Cambridge Isotope Laboratories) in M9 minimal media. Lyophilised IGF2 (Novozymes Biopharma, AU) was added to NMR samples of ^{15}N-single (dynamics studies) and ^{15}N, ^{13}C-double labeled domain 11^{AB5} mutant (structural studies) in a 1:1 ratio at pH 4, which provided optimum stability for the complex. Structural calculations of the domain 11^{AB5} both free and complexed with IGF2 were as described for the domain 11^{AB3}:IGF2 complex (5). Validation was performed using the iCing online server (version r76; (https://nmr.cmbi.ru.nl/icing/iCing.html). The ensembles of NMR structures and associated NMR chemical shifts have been deposited with the protein database and BioMagResBank with the following accession codes: Free domain 11^{AB5} 2M6T and IGF2 bound domain 11^{AB5} 2M68.

^{15}N relaxation studies of domain11^{WT} and domain 11^{AB5}.

NOE, ^{15}N-T_1 and ^{15}N-T_2 NMR relaxation data were acquired at two fields (600 and 900 MHz) at 25 ºC on a Varian VNMRS or INOVA spectrometer respectively, both fitted with cryogenically cooled probe heads for domain 11^{WT} and domain 11^{AB5}. T_1 delays of 0.01, 0.02*, 0.03, 0.05, 0.07, 0.10, 0.15, 0.30, 0.40, 0.6, 0.8, 1.0, 1.5, 2.0** and 2.5** s (*600 MHz only, **900 MHz only) and T_2 delays of 0.1-2.1 s with intervals of 0.2 s were gathered using BioPack pulse sequences provided by VnmrJ 2.2. T_1 and T_2 experiments were run in duplicate. Automatic spectral compression (ASCOM) (6) was applied to increase signal:noise without increasing the length of the experiments. NOE ratios were calculated between peak heights in HSQC spectra recorded both with and without NOE proton saturation. At the beginning of each transient a 5 s ultra low power pulse was employed for those spectra without NOE and a 3 s proton saturation pulse with a 2 s delay (total 5 s) for those experiments with NOE. Data collected was processed (NMRPipe, (7)) and assigned (Analysis 2.1.5 (8)) before being exported for curve fitting and model free analysis (relax 1.3.9 (9, 10). Relaxation rates were calculated by an eleven parameter grid search, followed by simplex minimisation curve fitting to a two-parameter (I_0, R_x) exponential decay. Monte Carlo error analysis used the standard deviation of simulated curves back calculated from fitted parameters across replicated spectra, or from the Sparky 3.113 (Goddard T. D.), with RMSD calculation of noise in a peak-free region of non-replicated spectra. Relaxation datasets recorded at 600 and 900 MHz were validated using self-consistency tests within relax (See Supplementary Figure 3). Extended model free analysis (11-13) was performed within relax 1.3.9 which involved AIC model-free model selection (14) and elimination of failed models followed by Monte Carlo simulations (15). After initial calculation of local τ_m values, isotropic (sphere) and anisotropic (prolate/oblate spheroids, ellipsoid) diffusion models were optimised followed by AIC selection of the best model.

Expression, purification and crystal structure of domain 11$^{AB5\text{-}RHH}$

Domain 11$^{AB5\ Q1569R\ P1597H\ S1602H}$ pET 26a plasmid was generated and refolded as above. Protein concentration of 2.5 mg/ml (160 μM) was used to screen for crystals in a 96 well format using Phoenix Liquid Handling System (Art Robbins Instruments) with multiple commercial buffer screens. Crystals corresponding to data collection were formed in JCSG-*plus*^{TM} screen, well G11 (2.0 M Ammonium Sulphate, 0.1M Bis-(2-hydroxyethyl)imino-tris(hydroxymethyl methane, pH 7.0), to generate orthorhombic crystals (width:5 μm, length: 40 μM). Data collection carried out using PILATUS detector at station I04 of the Diamond synchrotron revealed that crystals were orthorhombic and belonged to space group $P2_12_12_1$.

Diffraction images were processed using automated Xia 2 pipeline at Diamond Light Source, UK. Auto processed data generated from Xia 2 yielded intensity data to medium resolution (for data statistics, molecular replacement and refinement details, see Supplementary Table 8).

SPR validation and binding analysis.

Kinetic analysis of mutants by SPR was conducted using a Biacore T200 biosensor at 25 ºC in either HBS-EP or MES-EP (20 mM MES pH 6.5, 150 mM NaCl, 3 mM EDTA, and 0.005% (v/v) surfactant P20) at a flow-rate of 40 μL min^{-1}. For analysis of the domain 11:IGF2 interaction, approximately 20-50 RU of biotinylated IGF2 (GroPep, Aus) or IGF1 were immobilized on a CM5 chip by affinity capture to streptavidin and domain 11 at concentrations ranging from 64 nM to 0.125 nM. For the analysis of the Fc-domain 11^{AB5} interaction, approximately 50 RU of a biotinylated IGF2:biotinylated ubiquitin mixture (1:8 proportion) was immobilised on the CM5-streptavidin chip. Alternatively, Fc-domain 11^{AB5} and Fc-domain $11^{AB5\text{-}RHH}$ (or Fc-domain 11^{I1572A} as control) were immobilised by antibody capture (GE Healthcare) with and without human IgG1 Fc fragment (Bethyl Laboratories, UK) in the control flow cell. Concentrations of $IGF2^{1\text{-}67}$, $proIGF2^{1\text{-}104}$, $proIGF2^{1\text{-}156}$ and IGF1 ranging from 0.125 to 32 nM were injected. T200 BiaEvaluation software (GE Healthcare) was used to fit the resulting curves according to a 1:1 binding model. Thermodynamic data was performed across the temperature range of 10 ºC to 30 ºC, van't Hoff and Erying plots generated and ΔH, $-T\Delta S$ and ΔG values determined as previously described (4).

Fc-fusion protein expression and purification.

IgG1 Fc-domain 11^{AB5} fusion protein was first cloned as previously described (16), produced by transient transfection of HEK293T cells with pHLSec-AB5-hIgG1 and utilising Corning Cellstack10's to scale-up cell culture (Supplementary Fig. 7). Bulk production was performed in Chinese Hamster Ovary cells GS knockout (CHOK1SV GS-KO cells, Lonza Biologics, PLC). Prior evaluation of *in silico* antigenicity utilised Epibase™ HLA class II allotypes against amino acid sequences of Fc constructs, including linkers, showed that domain 11^{AB5} was less antigenic domain 11^{AB5} and so was selected for bulk production. Bulk cultures from stable selected CHOK1SV GS-KO resulted in protein A affinity purified yields in the range of 27 to 187 mg/L. SDS-Page and size exclusion HPLC confirmed a single Fc product with endotoxin levels < 0.74 EU/mg. For Fc-domain11^{AB5}:IGF2 co-precipitation, an aliquot of 400 μL of media of cells treated with Fc-domain11^{AB5} or Fc-domain11^{I1572A} or a previously incubated mixture of recombinant IGF2 and Fc-domain11^{AB5} was incubated overnight at 4 ºC with Protein A/G Agarose (Immuno-precipitation, Protein A, Roche).

Agarose beads were washed according to the manufacturer´s instructions, resuspended in 2 × sample buffer, boiled and the supernatant used for western blot analysis as described previously.

IGF2 induced hypoglycaemia.

All animal experiments were approved by the animal use ethical committee of Oxford University and fully complied with UK Home Office guidelines (PPL 30/2695). Wild-type C57BL/6J and 129S2 mice were anaesthetised using a rising concentration of isoflurane in oxygenated air and maintained using 2% isoflurane with 2 L min^{-1} oxygen. Blood glucose measurements were taken using an Advantage diabetic glucose meter (Roche Diagnostics, UK), blood glucose levels were allowed to normalise for 30 min prior to taking a baseline blood glucose reading, then either 1 mg kg^{-1} IGF2 alone (R&D systems), or 1 mg kg^{-1} IGF2 premixed with 10.9-11.7 mg kg^{-1} Fc-domain 11^{AB5}, Fc-domain $11^{AB5\text{-}RHH}$ or Fc-domain 11^{I1572A} (for a 1:1 molar ratio) were injected. Blood glucose measurements were taken every 10 min for the following hour. For the Fc-domain 11^{AB5} mouse preloading experiment, 11.7 mg kg^{-1} Fc-domain 11^{AB5} or PBS, as negative control, were injected immediately after induction of anaesthesia, 30 mins later a baseline blood glucose determined and IGF2 (1:1 molar ratio) injected, readings were then taken every 10 min for the following hour. For the IGF2 preloading experiment, 1 mg kg^{-1} IGF2 was injected immediately after induction of anaesthesia, 30 min later a baseline blood glucose reading was determined and either 11.7 mg kg^{-1} Fc-domain 11^{AB5} (1:1 molar ratio) or PBS were injected, readings were then taken every 10 min for the following hour.

IGF2 dependent human cancer cell xenografts.

IGF2 sensitive autocrine cell lines SKNMC-$IGF2^{1\text{-}67}$, SKNMC-FLAG-$IGF2^{1\text{-}67}$, SKNMC-$IGF2^{1\text{-}104}$and SKNMC-FLAG-$IGF2^{1\text{-}104}$ were generated from the Ewing sarcoma cell line SKNMC by retroviral infection with constructs containing IRES-luciferase. Constructs were generated by cloning into the the BglII and EcoRI of the pMSCV retroviral expression vector. The constructs contained the IGF2 secretion signal, followed by the IGF2 (e.g. with and without the FLAG-tag and R67STOP/ R104STOP truncated coding sequence) IRES and the Luc2 gene. Retrovirus were packaged by transfection of constructs with TransIT-X2 (Mirus) into Phoenix cells, medium was collected after 48 h, filtered (0.45 μm) and replaced the medium of SKNMC cells with polybrene (6-8 μg/ml). After 48hr, the transfection medium was removed from target cells and transduced cells were selected with fresh medium containing 2 μg/ml of puromycin.

Autocrine growth and signaling function of SKNMC lines, including inhibition by IGF2-TRAP, were confirmed in culture using an MTS reagent (Cell Titer 96 Aqueous One Solution cell proliferation assay, Promega). For xenografts, approximately $5x10^6$ SKNMC-$IGF2^{1-67}$ cells harvested during the growth phase of sub-confluent cell cultures were then mixed with matrigel (50:50 v/v, 100µl/ mouse) and injected (sub-cutaneous) into CD-1® (Crl:NU-1 *Foxn1^{nu}*) female mice, 24 hours after establishing ALZET® mini-pump placement delivering a single concentration 40mg kg^{-1} per week either IGF2-TRAP or PBS vehicle control. Tumour growth was monitored with both caliper measurements and bioluminescence. The latter was performed with a Bruker Xtreme 2D optical system 15 minutes (10 second exposure) following luciferin injection (100 µl per 20g body weight). Tumours were excised at 28 days or when size was > 1000 mm^3. Fixation in formalin, embedding in paraffin and Hematoxylin and Eosin (H&E) staining was done using standard protocols. Apoptotic area was quantified using automated image analysis of brightfield images in ImageJ, applying H AEC color de-convolution filter and automated thresholding. Serum was also collected at the time of culling.

IGF2-TRAP synthetic lethality screens.

For synthetic lethality screens (Target Discovery Institute, University of Oxford), either 2,500 or 5,000 SKNMC-$IGF2^{1-104}$-IRES-luc cells were plated in medium containing 0.5% charcoal stripped FBS in 384 well plates and the IGF2-TRAP was added 24h later to half of the plates. Using automated dispensing, 307 oncology drugs that have been tested in man (TDI expanded OncologyDrug library) were added to cells 3hrs later at 10 µM, 1 µM, 100 nM and 10 nM, in the presence and absence of 250 nM IGF2-TRAP. After 48hr of incubation, resazurin was added to a final concentration of 10 µg/mL and the fluorescence (Ex: 560 nm, Em: 590 nm) recorded at 2hr. Normalisation was performed with untreated cells and cells treated with the fixed concentration of IGF2-TRAP alone. Pearson's product-moment correlation coefficient (*r*) was calculated to determine replicate correlation. For hit selection, we used a bespoke 'interaction score' metric (a variation on standard Δz-score analysis) followed by hit selection using the robust non-parametric Rank Product method (17).

The interaction score was calculated where v was the viability readout for the drug condition, $\bar{x}$ was the sample mean of the negative controls and s was the sample standard deviation of the negative controls, z was calculated as $z_{treatment} = (v - \bar{x})/s$. z_{no_IT} was calculated from PBS control plates, comparing 'drug *d*' to same- plate negative controls 'cells + DMSO'. z_{IT} was calculated from treated plates, comparing 'drug *d* + IGF2-TRAP (IT)' to same-plate negative controls 'cells + IT + DMSO'. The magnitude of these z-scores reflects

the degree of inhibition of cell viability by 'drug d ± IT' relative to negative controls, also taking into account well-to-well variation. The interaction score is defined as $z_{no\_LT} - z_{IT}$. Thus, a more negative interaction score represents a greater synergistic lethal effect: the inhibitory effect of the drug is potentiated by IT. Interaction scores were calculated for each replicate $r_{d,1}/r_{d,2}/r_{d,3}$ independently.

The rank product was calculated in each data set (16 data sets: cell density x drug library concentration x rezasurin incubation time) each replicate for each drug was ranked independently, with 1st rank assigned to the most negative interaction score. Using these ranks, a Rank Product $RP(d)$ for drug d in each data set was calculated, where $r_{d,i}$ is the rank of the ith replicate for drug d:

$$RP(d) = \sqrt[3]{(r_{d,1})(r_{d,2})(r_{d,3})}$$

In a permutation test, the RP for each drug is compared to a simulated data set giving an estimated p-value of false positive. These p values, in the context of our experiment, informed us about effect size, direction and reliability, and allowed us to control the false discovery rate in contrast to many other hit selection protocols[43]. Hits were selected where $p<0.05$.

For synthetic lethality hit validation, fresh stocks of the two drugs selected as hits were obtained (pictilisib, SYNkinase and PF-04691502, Cayman Chemical). The interaction of the selected drugs and IGF2-TRAP were evaluated by Q-value calculated by the formula $Q=F_{D+T}/F_D+(1-F_D)F_T$, where F_{D+T} represents the fraction affected by treatment with the drug plus IGF2-TRAP compared with the untreated control group, F_D represents the fraction affected by the drug alone, and F_T represents the fraction affected by the IGF2-TRAP alone. A value of $Q>1.15$ indicates a synergistic effect, $Q<0.85$ indicates an antagonistic effect, and Q between 0.85 and 1.15 indicates an additive effect (18).

In addition to the screen protocol, we used cell lines SKNMC-IGF2$^{1\text{-}67}$-IRES-luc and Ewing's line ST-A-ET1 (ATCC) and a tailored dose range for each drug (Supplementary Figure 11). Standard curve fit used: log (inhibitor) vs. response (three parameters) with constant slope = -1.0. R^2 value for goodness-of-fit and IC_{50} values with 95% confidence intervals. Error bars indicate s.e.m. Drug concentration of zero is shown on log dose plots as ~2 logs below the next lowest dose. Prism 6 software was used for all statistical analysis apart from hit selection.

Cell assays, protein extraction and Immunoblotting.

For analysis of IGF2 secretion by the different cell lines, cells were plated in 6-well plates in serum-free medium at a concentration of approximately 10^6 cells per well. The medium was collected after 48 hr, spun at 4000g for 20 min and the supernatant filtered through a 0.22 μm filter. For signaling analysis, cells were plated in 6-well plates at a concentration of approximately 3-6x10^5 cells per well in serum-free medium. After 24 hr, Fc-domain$11^{AB5/AB5-RHH}$ (or PBS and Fc-domain 11^{I1572} as negative controls) were added to each well to the corresponding concentration and the cells were incubated for further 48-72 hr. Cells were washed with PBS and lysed on ice in RIPA buffer (50mM Tris HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) containing protease and phosphatase inhibitors (Halt cocktail, Pierce). Lysates were cleared by centrifugation at 4°C. Protein levels were quantified using the CB-X assay (G-Biosciences) as per manufacturer's instructions. SDS-PAGE gel electrophoresis and protein transfer of either 20 μg of protein or 5 μL of serum were done following standard protocols. Antibodies used were anti hIGF2 (R&D Systems AF-292, 2 μg mL^{-1}), anti $pAkt^{S473}$ (Cell Signaling 4060, 1:1000), anti β-actin (AbCam 8227, 1:20000), anti $pIGF1R^{Y1135/1136}$/$pIR^{Y1150/1151}$ (Cell Signaling 3024, 1:1000), CaptureSelect Biotin Anti-hIgG-Fc (Life Technologies, 1:1000), anti mIGF1 (R&D Systems AF-791, 1 μg mL^{-1}), anti mIGFBP2 (Millipore 06-107, 1/1000 dilution), anti mIGFBP3 (R&D Systems MAB775, 1 μg mL^{-1}), anti mIGFBP5 (R&D Systems AF578, 0.1 μg mL^{-1}), anti mIGFBP6 (R&D Systems AF776, 0.1 μg mL^{-1}), anti mGH (R&D Systems BAF1566, 0.5 μg mL^{-1}), anti hIGFBP3 (Santa Cruz Biotechnology Sc-9028, 1 μg mL^{-1}), anti hALS (Novus Biologicals NBP1-89118, 1 μg mL^{-1}). Secondary antibodies were from Dako, used at 1:2000, except Streptavidin HRP at 1:4000 (Thermo Scientific Pierce). Densitometry was performed using ImageJ (NIH) and was normalized to actin.

Supplementary Tables

Supplmentary Table 1

Loop libraries. Mutated residues in each loop are shown in bold, Sequences of the forward and reverse oligonucleotide primer pairs used for site-directed mutagenesis to generate domain 11 loop specific libraries are shown. One oligo in each pair was 5'-phosphorylated (indicated with *). XXX indicates an equal amount of codons encoding each one of the 19 possible amino acids (Cys was excluded).

Mutant loop Library	Sequence	Oligonucleotide Sequence (5' to 3')	Number of possible variants
AB	**YSKKGL**	**NNK NNK NNK** GTT TAC ATG AGC ATC TGT GGG GAG *CTT **MNN MNN** AGC AGC TGT GAA TCC CGC CC	3.2 x10^6
HI	**DKQT**	***XXX XXX** TGC ACT CTC TTC TTC TCC TGG **XXX XXX** CAG GGA GAT GAG CAT GGG	1.3 x10^5

Supplementary Table 2. **Affinity kinetics of AB loop mutagenesis to identify domain 11^{AB5} and site directed mutants following structural predictions.** SPR binding analysis of recombinant M6P/IGF2R domain 11 mutants interaction with IGF2. k_{on} indicates association (on) rate; k_{off}, dissociation (off) rate; *RU*, response units. Binding to IGF2 assessed with BIAcore T200 using purified biotinylated IGF2 and IGF1 on a streptavidin biosensor chip. Recombinant proteins were expressed either in *E. coli* and/or *P. pastoris*. All the mutants in the FG and HI loops are on the domain 11^{AB5} background for the AB loop unless otherwise stated. All experiments were carried out at 25 °C in HBS-EP buffer (pH 7.4) and repeated 2-5 times on two independent BIAcore chips. nd = not determined. Data for K_D, k_{on} and k_{off} are expressed as means. * AB3 background for the AB loop.

Loop	Mutant Domain 11	Kinetics				Steady state	
		K_D (nM)	**k_{on} (x10^6 M^{-1} s^{-1})**	**k_{off} (s^{-1})**	**Chi^2 (RU^2)**	**K_D (nM)**	**Chi^2 (RU^2)**
AB	WT	46±1.8	5.0	0.23	0.33	64±1.0	0.24
	AB3	15.3±2.0	2.76	0.044	0.41	15.5±3.3	0.77
	AB5	**5.07±0.87**	**4.21**	**0.0208**	**0.87**	**6.44±1.06**	**0.14**
	W1546F	4.35±0.49	4.5	0.0180	0.767	5.00±0.93	0.129
FG	P1599A*	>70	Nd	>0.1	-	>70	-
	P1599V*	>100	Nd	>0.1	-	>100	-
	P1599G*	>200	Nd	>0.2	-	>200	-
	S1600A	89.10±5.1	2.39	0.2126	nd	84.65±15.20	0.04
	S1600V	>200	Nd	>0.2	nd	>200	-
	S1600L	>200	Nd	>0.2	nd	>200	-
	S1600Y	>200	Nd	>0.2	nd	>200	-
	S1600N	24.44±8.70	3.95	0.0910	0.15	29.67±9.54	0.01
	K1601A	6.80±0.65	3.26	0.0220	0.51	8.51±0.39	0.04
	K1601R	4.75±0.08	5.28	0.0251	0.37	5.78±0.18	0.07
	S1602A	16.48±3.83	1.57	0.0250	0.63	19.92±3.77	0.17
	S1602H	3.19±1.00	3.61	0.0112	1.24	5.44±1.41	0.96
	S1602N	4.75±0.52	2.38	0.0112	0.56	7.93±2.29	0.46
	S1602Q	3.61±1.19	5.88	0.0191	0.57	5.11±0.73	0.13
	G1603A	8.69±1.48	2.14	0.0183	0.45	12.09±0.69	0.16
	G1603K	2.73±0.95	4.66	0.0121	1.18	4.54±0.91	1.18
	L1604A	48.34±8.4	0.79	0.0397	0.26	59.96±19.8	0.01
HI	K1631A	>200	-	-	-	>200	-
	K1631R	12.29±2.81	7.18	0.0867	0.23	13.51±2.65	0.08
	K1631E	>100	-	>0.1	-	>100	-
	K1631W	2.94±0.37	1.18	0.0034	1.47	12.16±2.35	3.16
	Q1632A	3.76±0.54	4.50	0.0167	0.57	4.84±0.22	0.12
	Q1632K	11.07±2.10	1.96	0.0212	0.34	12.84±0.63	0.05
	Q1632E	39.83±7.10	0.52	0.0202	2.11	52.41±4.29	0.06
	Q1632L	5.75 ±1.05	3.18	0.0175	1.17	9.34±1.69	2.00
	T1633A	26.41±9.49	0.80	0.0196	0.21	53.55±10.0	0.19

Supplementary Table 3.

Affinity kinetics of Domain 11^{AB3} spontaneous mutants and HI Loop mutagenesis library directed screen.

Mutant Domain 11	Kinetics				Steady state	
	K_D (nM)	k_{on} (x10^{6}M^{-1} s^{-1})	k_{off} (s^{-1})	Chi^2 (RU^2)	K_D (nM)	Chi^2 (RU^2)
WT	46±1.8	5.0	0.23	0.33	64±1.0	0.24
AB3	15.3±2.0	2.76	0.044	0.41	15.5±3.3	0.77
AB3 + 62B (FG-P1597H)	1.23±0.32	10.6	**0.013**	1.62	1.76±0.47	1.43
AB3 + 73E (CD-Q1569R)	2.30±0.27	13.6	**0.031**	1.54	3.15±0.41	0.25
WT + 23D (HI-HFQS)*	176	0.25	**0.065**	0.56	220.5	0.45
WT + 26C (HI-DMQT)*	130	1.89	0.24	0.17	127	0.19
WT + 27G (HI-NMAL)*	195	2.0	0.25	0.19	194	0.13
WT+ 312G (HI-VKLM)*	119	1.07	0.15	0.27	135	0.45
AB5 + 51F (HI-NRQS)*	13.5	7.4	0.10	0.31	6.94	0.26

*Protein quality poor based on absorption spectra

Supplementary Table 4
NMR structural data for domain11^{WT}, domain 11^{AB5} and domain 11^{AB5}:IGF2 complex
Restraints, energies, RMSD values and number of NOE violations for the ensemble of twenty structures for each protein. All values calculated by ARIA(19) , with the exception of NOE restraint distribution, which was calculated using the countNOE.py script provided by Benjamin Bardiaux (Institut Pasteur, Paris). *RMSD to mean values were calculated for residues in the secondary structure.

For ensembles of 20 structures	**Domain 11^{WT}**	**Domain 11^{AB5}**	**Domain 11^{AB5}:IGF2 Complex**
NOE restraints			
Total NOE restraints	2200	2168	3176
Intraresidue NOEs	800	871	1242
Sequential/short/medium NOEs (residue *i* to *i+1, 2, 3, 4, 5*)	746	718	1158
Long-range NOEs	654	446	776
Dihedral angle restraints			
Backbone φ and ψ	206	211	249
Hydrogen bond restraints	60	48	52
Energies (kcal mol^{-1})			
E_{NOE}	3.3 ± 0.4	4.2 ± 1.8	7.0 ± 0.6
E_{VDW}	-567 ± 21	-1187 ± 13	-1226 ± 12
E_{ELEC}	-5531 ± 74	-5117 ± 98	-5064 ± 69
RMSD from ideal geometry			
Bonds (Å)	0.0048 ± 0.0001	0.0101 ± 0.0002	0.0105 ± 0.0002
Angles (°)	0.58 ± 0.01	1.03 ± 0.03	1.07 ± 0.02
Improper (°)	0.57 ± 0.02	1.26 ± 0.07	1.36 ± 0.07
RMSD to mean for backbone atoms*	0.54 ± 0.12	0.55 ± 0.10	0.27 ± 0.04
RMSD to mean for heavy atoms*	1.52 ± 0.22	0.99 ±0.10	0.56 ± 0.04
Average number of NOE violations			
>0.1 Å (per structure)	121	1	0.1
>0.3 Å (per structure)	11	0	0
>0.5 Å (per structure)	0	0	0

Supplementary Table 5

Structure quality statistics for domain11^{WT}, domain 11^{AB5} and domain 11^{AB5}:IGF2 complex

Ramachandran plot regions and iCing Z-scores for the lowest energy structure and ensemble for domain 11^{WT}, domain 11^{AB5} and its complex with IGF2. Structure Z-scores above zero are better than average. RMS Z-scores should be close to 1. Human ensemble results were recalculated using NRG-CING (20). Best refers to the best statistics selected for one model within the ensemble.

	Domain 11^{WT}		**Domain 11^{AB5}**		**Domain 11^{AB5}:IGF2 Complex**	
For ensembles of 20 structures	*Best*	*Ensemble*	*Best*	*Ensemble*	*Best*	*Ensemble*
Ramachandran plot						
Most favoured (%)	78.3	77.7	80.5	80.6	76.3	77.0
Additionally allowed (%)	17.4	17.9	18.6	18.2	22.0	22.1
Generously allowed (%)	3.4	3.5	0.8	0.6	1.7	0.8
Disallowed (%)	0.9	0.9	0.0	0.6	0.0	0.1
Structure Z-scores						
1st generation packing	-0.54	1.76 ± 0.62	-0.51	2.24 ± 0.81	-0.54	2.23 ± 0.59
2nd generation packing	-1.63	7.59 ± 2.20	-0.79	7.41 ± 2.15	0.95	7.41 ± 2.03
Ramachandran plot	-5.04	-4.24 ± 0.23	-2.85	-2.87 ± 0.43	-3.96	-3.92 ± 0.33
$\chi_1 - \chi_2$ rotamer normality	-3.45	-3.70 ± 0.44	-0.74	-1.38 ± 0.55	-2.67	-2.62 ± 0.40
Backbone conformation	-3.47	1.56 ± 0.22	-0.61	-0.94 ± 0.29	-1.46	1.47 ± 0.15
RMS Z-scores						
Bond lengths	1.28	1.28 ± 0.02	1.11	1.11 ± 0.01	1.12	1.12 ± 0.01
Bond angles	0.64	0.64 ± 0.02	0.46	0.46 ± 0.01	0.46	0.47 ± 0.01
ω angles	0.17	0.91 ± 0.05	0.56	0.58 ± 0.04	0.67	0.63 ± 0.03
Side-chain planarity	0.85	0.85 ± 0.10	0.63	0.62 ± 0.09	0.71	0.69 ± 0.10
Improper dihedrals	0.79	1.07 ± 0.05	0.78	0.78 ± 0.03	0.78	0.81 ± 0.02
Inside/outside distribution	0.95	0.94 ± 0.01	0.94	0.96 ± 0.02	0.97	0.97 ± 0.01

Supplementary Table 6.

Affinity kinetics of combined CD, FG and HI, loop combined mutants with AB loop backgrounds domain 11^{AB3} and domain 11^{AB5}. Measurements were carried out at 25°C in HBS-EP buffer (pH 7.4) (nd= not determined, np= no protein). Data for K_D expressed as mean of each experiment, and means of k_{on} and k_{off} are shown.

Domain 11 Mutant	*Kinetics*			*Steady State*	Number of mutations
	K_D (nM)	k_{on} ($x10^6 M^{-1} s^{-1}$)	k_{off} (s^{-1})	K_D (nM)	
AB3	15.3±2.0	2.76	0.044	15.5±3.3	3
AB3 Q1569R	2.30±0.27	13.6	0.031	3.15±0.41	4
AB3 P1597H	1.23±0.32	10.6	0.013	1.76±0.47	4
AB3 P1597K	2.78±0.61	10.5	0.0287	3.59±0.77	4
AB3 Q1569R P1597H	0.90±0.29	18.9	0.0153	1.73±0.4	5
AB3 Q1569R P1597K	1.34±0.47	20.0	0.0241	2.27±1.04	5
AB3 P1597H K1631W	17.55±1.88	19.7	0.3478	18.67±2.23	5
AB3 P1597H HI^{23D}	nd	nd	0.012	nd	8
AB3 P1597K S1602H	3.73	2.74	0.0102	5.332	5
AB3 Q1569R P1597H S1602H	0.74±0.22	14.1	0.010	1.53±0.46	6
AB3 Q1569R P1597K S1602H	0.79±0.25	20.1	0.016	1.39±0.48	6
AB3 Q1569R P1597H K1631W	np	np	np	np	6
AB3 Q1569R P1597K K1631W	np	np	np	np	6
AB3 Q1569R P1597H S1602H G1603K	0.62±0.03	17.5	0.0108	1.31±0.16	7
AB3 Q1569R P1597K S1602H G1603K	0.78±0.22	20.8	0.0171	1.62±0.41	7
AB3 Q1569R P1597H S1602H K1631W	np	np	np	np	7
AB5	5.07±0.87	4.21	0.0208	6.44±1.06	5
AB5 Q1569R	2.24±0.44	8.04	0.0175	3.22±0.55	6
AB5 P1597H	1.90±0.60	3.68	0.0070	2.79±0.52	6
AB5 P1597A	9.07±1.08	1.90	0.0174	12.44±0.69	6
AB5 P1597K	1.94±0.39	9.54	0.0120	2.21±0.14	6
AB5 P1597L	3.86±1.68	4.18	0.0150	4.85±1.82	6
AB5 P1597N	4.13±1.74	4.72	0.0180	5.36±2.24	6
AB5 P1597Q	3.83±2.42	6.53	0.0203	4.51±2.66	6
AB5 P1597S	4.19±1.80	5.00	0.0193	4.90±1.88	6
AB5 HI^{23D}	np	np	0.012	np	9
AB5 Q1569R P1597H	1.95±0.60	3.2	0.0058	5.40±0.98	7
AB5 Q1569R P1597K	0.95	9.59	0.0091	1.69	7
AB5 P1597H S1602H	1.17±0.2	3.50	0.0046	2.93±0.47	7

AB5 P1597K S1602H	4.62	2.09	0.0096	7.85	7
AB5 P1597H $HI^{2\text{-}3D}$	np	np	np	np	10
AB5 Q1569R P1597H S1602H	0.65±0.11	5.90	0.0038	2.0±0.50	8
AB5 Q1569R P1597K S1602H	0.92±0.13	7.87	0.0073	2.11±0.55	8
AB5 Q1569R P1597H G1603K	0.87±0.09	6.67	0.0058	2.01±0.30	8
AB5 Q1569R P1597H Q1632L	np	np	np	np	8
AB5 Q1569R P1597H K1631W	1.02±0.06	11.20	0.0115	1.95±0.47	8
AB5 Q1569R P1597K K1631W	np	np	np	np	8
AB5 Q1569R P1597H HI2-3D	np	np	np	np	11
AB5 Q1569R P1597H S1602H G1603K	np	np	0.005	np	9
AB5 Q1569R P1597K S1602H G1603K	np	np	0.008	np	9
AB5 Q1569R P1597H S1602H K1631W	np	np	np	np	9
AB5 Q1569R P1597K S1602H K1631W	np	np	np	np	9
AB5 Q1569R P1597K S1602H Q1632L	np	np	np	np	9
AB5 Q1569R P1597H S1602H G1603K K1631W	np	np	np	np	10
AB5 Q1569R P1597K S1602H G1603K K1631W	np	np	np	np	10

Supplementary Table 7.

pH dependent affinity kinetics of combined CD, FG and HI, loop mutants with AB loop backgrounds domain 11^{AB3} and domain 11^{AB5} (pH 6.5) compared to Table 6 (pH 7.4). Measurements were carried out at 25°C in MES-EP buffer (20 mM MES pH 6.5, 150 mM NaCl, 3 mM EDTA, and 0.005% (v/v) surfactant P20). Data for K_D expressed as mean of each experiment, and means of k_{on} and k_{off} are shown.

Mutant	***Kinetics***		
	K_D **(nM)**	k_{on} **(x10^6 M^{-1}s^{-1})**	k_{off} **(s^{-1})**
WT	35.06±10.45	2.68	0.1441
AB3	11.85±0.20	3.43	0.0381
AB3 Q1569R P1597H	0.47±0.06	19.9	0.0091
AB3 Q1569R P1597H S1602H	0.40±0.07	16.1	0.0064
AB3 Q1569R P1597H S1602H G1603K	0.43±0.03	24.3	0.0087
AB5	4.38±0.78	4.17	0.0180
AB5 Q1569R P1597H	1.46±0.29	2.37	0.0034
AB5 Q1569R P1597H S1602H	0.71±0.10	4.41	0.0030
AB5 P1597H S1602H	0.62±0.34	5.32	0.0029
AB5 Q1569R P1597H G1603K	0.69±0.06	7.54	0.0051

Supplementary Table 8

Crystallographic Data Statistics for Domain 11 $^{AB5\text{-}RHH}$

Data set	Domain 11 $^{AB5\text{-}RHH}$ 2.8 Å
Space Group	$P2_12_12_1$
Wavelength (Å)	0.98
High resolution Range (Å)	34.9-2.8
Measurements	42539
Unique Reflections	6912
Completeness (%)	99.9 (100)[a]
$I/\sigma(I)$	6.4 (1.35)
R_{merge} (%)[b]	22.7 (46.4)
CC half	0.989
Wilson B factor	26.5
R factor (%)[c]	23.7
R_{free} (%)[d]	27.8
R.m.s.d. from ideal	
Bonds (Å)	0.017
Angles (°)	1.90
Average B factors (Å)	30.5
B factor r.m.s.d. (Å)	
main chain	1.32
side chain	1.65
Total number of non-hydrogen atoms	
protein	1018
water	20
other	2x SO_4^{2-} 1 x glycerol 1x 1,2-ethanediol

[a] Values in parentheses correspond to the highest resolution data shell (2.8 Å)

[b] $R_{merge} = \Sigma|I-\langle I\rangle| / \Sigma\langle I\rangle$

[c] R factor $= \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$

[d] R_{free} is as for R factor but calculated using a 5% test set of reflections excluded from the refinement.

Data collection carried out using PILATUS detector at station I04 of the Diamond synchrotron revealed that crystals were orthorhombic and belonged to space group $P2_12_12_1$. Diffraction images were processed using automated Xia 2 pipeline. Auto processed data generated from Xia 2 yielded intensity data to medium resolution. Processing was carried out using programs in the CCP4 suite (21). Phaser MR (22) was used in molecular replacement (MR) with a known structure (PDB:1GP0) for domain11^{WT} to generate a MR solution. The crystals contained a single domain 11$^{AB5\text{-}RHH}$ monomer per crystallographic asymmetric unit; the same as 1GP0. Structural refinement was then carried out using refmac5 (23). Further refinement was undertaken, with the rebuilding of the AB loop from WT to AB5, the introduction of residues Q1569R, P1597H and S1602H and other manual rebuilding performed using COOT (24). Structural model built was then validated using PROCHECK (25), MolProbity (26), and PDB_REDO server (27).

Supplementary Table 9

Affinity of IGF2 isoforms ($IGF2^{1-67}$, $IGF2^{1-104}$ and $IGF2^{1-156}$) and IGF1 for the Fc-domain 11^{AB5}, Fc-domain $11^{AB5\text{-}RHH}$ and Fc-domain 11^{I1572A} fusion proteins. SPR analysis of the interaction of recombinant IGF1 and IGF2 forms with Fc-domain 11 proteins immobilised by antibody capture (CM5 chip containing monoclonal mouse anti-human IgG antibody immobilised by amine coupling). Measurements were carried out at 25°C in HBS-EP buffer (pH 7.4). Example sensorgrams are shown in Fig. 5A.

IGF isoform	***Kinetics***		
	K_D **(nM)**	k_{on} **(x10^6 M^{-1}s^{-1})**	k_{off} **(s^{-1})**
Interaction with Fc-domain 11^{AB5}			
$IGF2^{1-67}$	3.89	6.47	0.025
$IGF2^{1-104}$	0.54	12.87	0.0069
$IGF2^{1-156}$	0.21*	4.19*	0.0009
IGF1	-	-	-
Interaction with Fc-domain $11^{AB5\text{-}RHH}$			
$IGF2^{1-67}$	0.54	2.60	0.0014
$IGF2^{1-104}$	0.39	1.29	0.0005
$IGF2^{1-156}$	1.13*	0.23*	0.0003
IGF1	-	-	-
Interaction with Fc-domain 11^{I1572A}			
$IGF2^{1-67}$	-	-	-
$IGF2^{1-104}$	4.41	0.52	0.002
$IGF2^{1-156}$	2.71*	0.29*	0.0008
IGF1	-	-	-

* Note, kinetic analysis may be inaccurate because of uncertainties in the concentration of recombinant $IGF2^{1-156}$ (Gropep). This consisted of a mixture of three isoforms, the full-length protein and truncated peptides (151 and 141 residues long).

Supplementary Figures

Background to Supplementary Figures S1-4

Relaxation data analysis for domain 11^{WT} and domain 11^{AB5}.

NOE, R_1 and R_2 relaxation datasets were recorded for domain 11^{WT} and domain 11^{AB5} and are shown in Supplementary Fig. 2a-f. NOE ratios, which are sensitive to fast picosecond motions, are shown at both 600 and 900 MHz for domain 11 and domain 11^{AB5} (Supplementary Fig. 1a, 1b). NOE values above 0.7 (indicating more ordered regions of protein structure) were generally located in β-strand secondary structure of the β-barrel in both proteins as previously observed for domain 11^{WT} (5). The AB, CD, FG and GH loop regions and termini had lower NOE ratios indicating the presence of faster ns-ps motions. The most flexible loops in domain 11^{WT} were the AB and GH loops with significantly reduced values seen in both. In domain 11^{AB5} however, the AB loop NOEs does not show the same degree of reduction. Flexibility in the isolated domain 11^{WT} AB and GH loops are might be expected due to removal of the neighbouring domain 13 that is observed to pack against it in the X-ray structure (28). The FG loop shows reduced NOE values in both proteins whereas the CD loop appears to show only marginally lower values. These differences are also reflected in the R_1 values with domain 11^{AB5} showing almost uniform values across the entire peptide compared to domain 11^{WT} which whose elevated R_1 values across the loops with reduced NOE values. With the exception of the FG loop, R_2 values across domain 11^{AB5} are reduced compared to the wild type human protein.

Validation of domain 11^{WT} and domain 11^{AB5} relaxation datasets.

R_1 and R_2 relaxation rates were validated by calculation of 900:600MHz ratios for each parameter (Supplementary Fig. 2). R_1 values were more consistent than R_2 values due better signal to noise in these experiments. $R_1(900)/R_1(600)$ and $R_2(900)/R_2(600)$ ratio averages (Supplementary Fig. 2a-d) and standard deviations were comparable between the two proteins. $R_1(900)/R_1(600)$ ratios were 0.65 ± 0.07 for human wild-type and 0.64 ± 0.05 for domain 11^{AB5}. $R_2(900)/R_2(600)$ ratios were 1.30 ± 0.22 for domain 11 and 1.17 ± 0.29 for domain 11^{AB5}. R_2/R_1 ratios were calculated at 600 (black) and 900 (white) MHz (Supplementary Fig. 2e-f) with lines marking the mean for each field. 900:600MHz ratios deviating form the average and yielding a non-linear correlation between R_2/R_1 and field strength, were excluded from subsequent Lipari-Szabo model-free calculations. An average of the three consistency test results is shown for domain 11^{WT} (Supplementary Fig. 2g) and domain 11^{AB5} (Supplementary Fig. 2h). For all tests, a parameter ratio 900/600 MHz close to 1 represented data which was consistent between fields. Results were more consistent for domain 11 (0.98 ± 0.21) versus 1.07 ± 0.16 for domain 11^{AB5}. $F\eta$, F_{R2} (29) and $J(0)$ (30)consistency tests provided with RELAX used NOE ratio, and R_1 and R_2 relaxation rates (31) are shown in Supplementary Fig. 3a-l. Distributions (panels d-f for domain 11 versus j-l for domain 11^{AB5}) for both proteins are within a range of values previously described (31-33). Any values that severely deviated from a 900:600MHz ratio of 1 were removed before input of parameters into relax for application of the Lipari-Szabo model-free formalism.

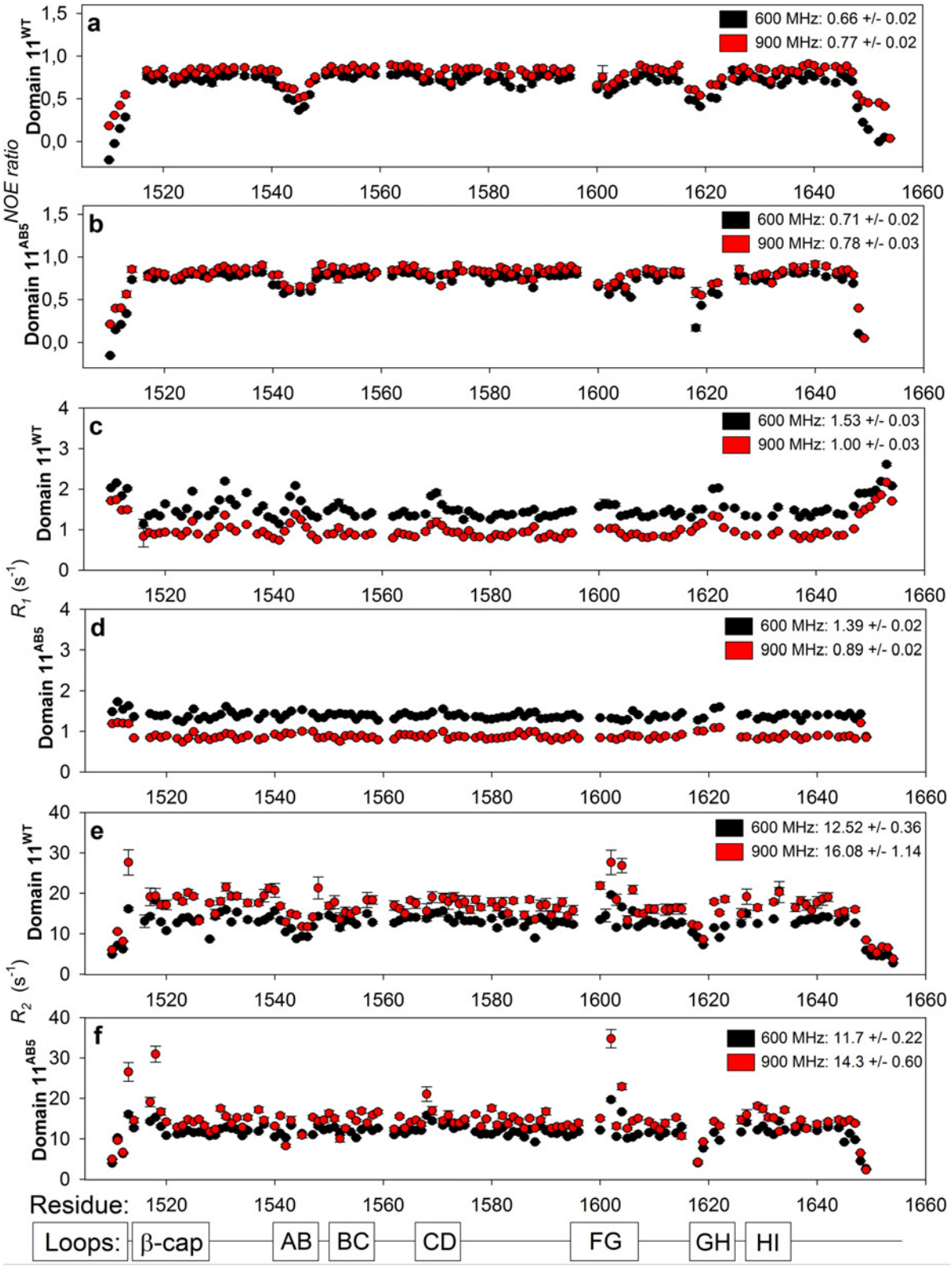
Lipari-Szabo model-free calculations for domain 11^{WT} and domain 11^{AB5}.

Lipari-Szabo parameters for isotropic (spherical) and anisotropic models (oblate, prolate and ellipsoid) were best fit in RELAX by oblate and fully anisotropic models for domain 11^{WT} and domain 11^{AB5} respectively. Residues that passed internal consistency tests were used in the calculation of the global rotational correlation time that was determined to be 7.62 ns for domain 11^{WT} and 8.19 ns for domain 11^{AB5}. This is consistent with a negligible effect on the oligomeric state of the protein and the overall tumbling rate in solution. Residues were fit to one of five models. These included model 1, where the backbone motion could be described by the generalized order parameter S^2 only, model 2 that incorporates S^2 and a contribution from faster picosecond motions characterized by a second correlation time τ_e (S^2 - τ_e) and models 3 and 4 that incorporate an additional slower (typically μs-ms) exchange contribution (R_{ex}) and described by S^2 - τ_e and S^2 - τ_e - R_{ex} respectively. Fits to these models are shown in Supplementary Fig. 4a-c. Domain 11^{WT} flexible AB, BC and CD loops, shown by S^2 values below 0.7. These loop regions are much more rigid in domain 11^{AB5}, consistent with NOE, R_1 and R_2 data. This suggests that the ring-stacking interactions of Y1542, W1546 and F1567 could reduce flexibility in the AB, BC and CD loops, stabilising the docking site of IGF2. Internal correlation times (τ_e) also reveal values of ~ 1ns that are present in the β-cap, pre-AB loop strand, AB, CD, FG and GH loops of domain 11^{WT} that are not observed in domain 11^{AB5}. Similarly greater contributions from slower exchange are observed in the β-cap, AB, BC, CD, FG and GH loops in domain 11^{WT}. In domain 11^{AB5} these again appear to be lost with the exception of the FG loop, which is unaffected.

Residues that could not be defined by these simpler models were also fitted to a two-time-scale model (34), an extension of the original Lipari-Szabo formalism that incorporates both fast picosecond (S^2_f, τ_f) and slow, sub-nanosecond (S^2_s, τ_s where S^2= $S^2_f S^2_f$) contributions to the overall correlation function for the amide bond vector. S^2_f, S^2_f, τ_f, τ_s, parameters are shown in Supplementary Fig. 4d-g. Residues 1539-1541 and 1547 in domain 11 display lower S^2_f values, mirroring the reduction in S^2 seen in the AB loop whereas residues 1544-1546 instead show lower S^2_s values which still indicates flexibility but on a slower timescale and may reflect their proximity, and packing, against the the more rigid CD loop. In domain 11^{AB5} both fast and slow motions are lost in the AB loop emphasizing the importance of the G1546Q mutation in stabilizing the binding site of this mutant.

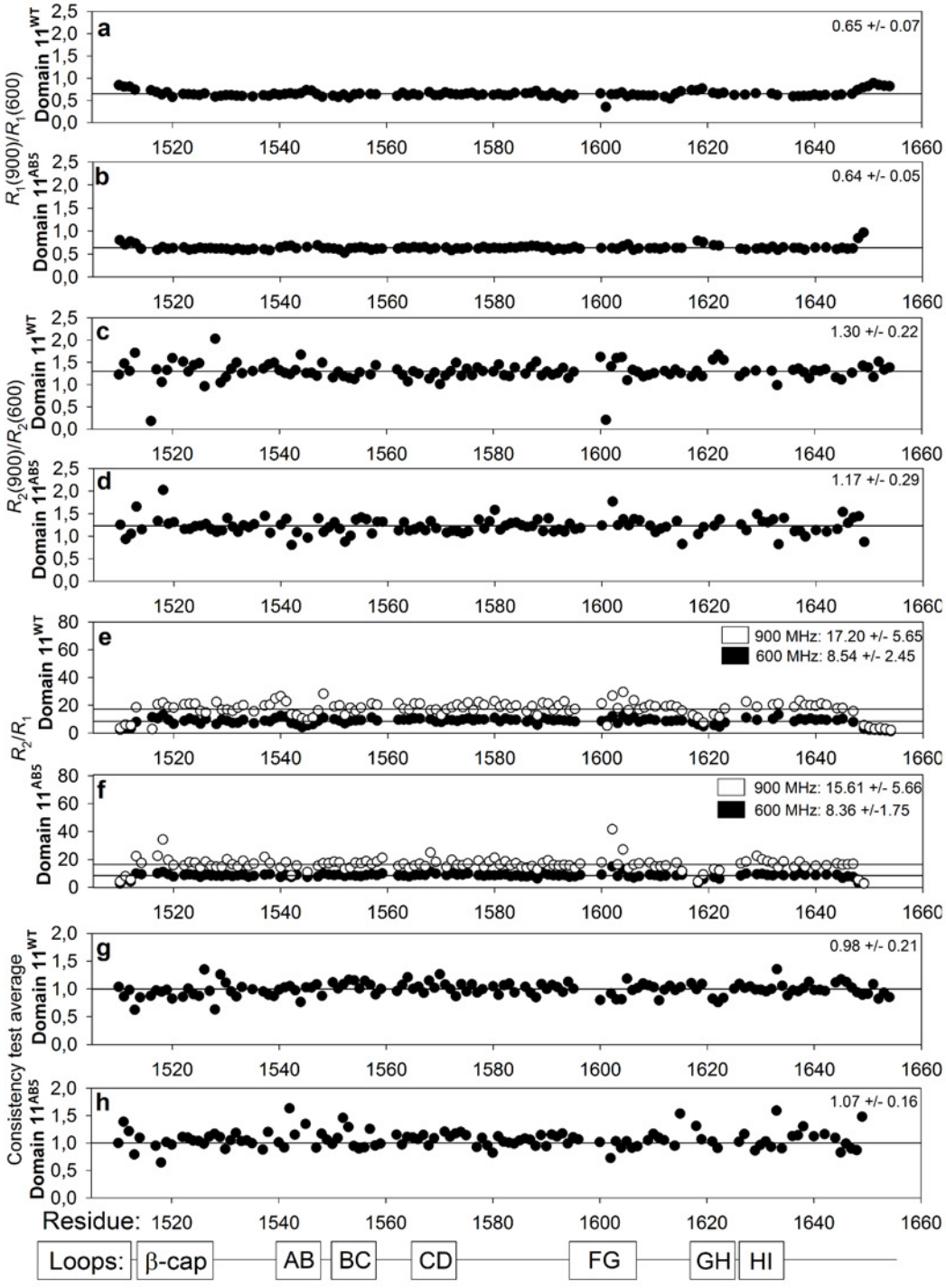
Supplementary Figure S1.

Relaxation data for human domain 11^{WT} and domain 11^{AB5}. Relaxation rates for 600 (black) and 900 and MHz magnetic field strengths are shown (red) are shown. (**a**) and (**b**) show NOE ratios for domain 11^{WT} and domain 11^{AB5} respectively. (**c**) and (**d**) show R_1 (s^{-1}) and (**e**) and (**f**) show R_2 (s^{-1}) for domain 11^{WT} and domain 11^{AB5} respectively.



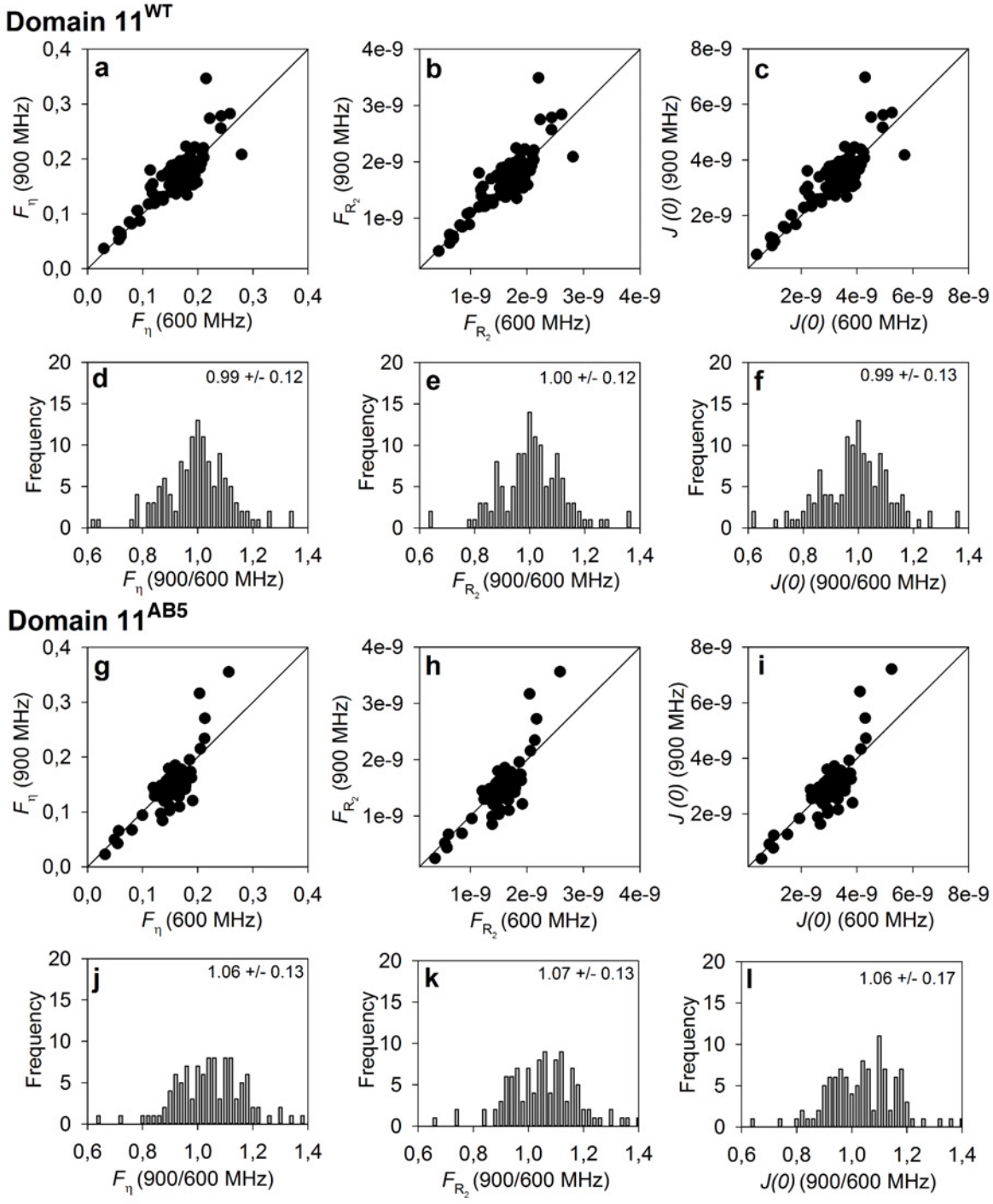
Supplementary Figure S2.

Consistency tests for human domain 11^{WT} and domain 11^{AB5} relaxation datasets. (**a**) and (**b**) R_1 and (**c**) and (**d**) R_2 (900/600 MHz), (**e**) and (**f**) R_2/R_1 values at two magnetic field strengths with black (600 MHz) and white (900 MHz) circles for domain 11^{WT} and domain 11^{AB5}. (**g**) and (**h**) Consistency test averages for human domain 11^{WT} and domain 11^{AB5} where values nearer 1 are more consistent between fields (y = 1 is plotted). Averages and standard deviations are shown in the top right corner of each graph (averages are also plotted in the top three panels).



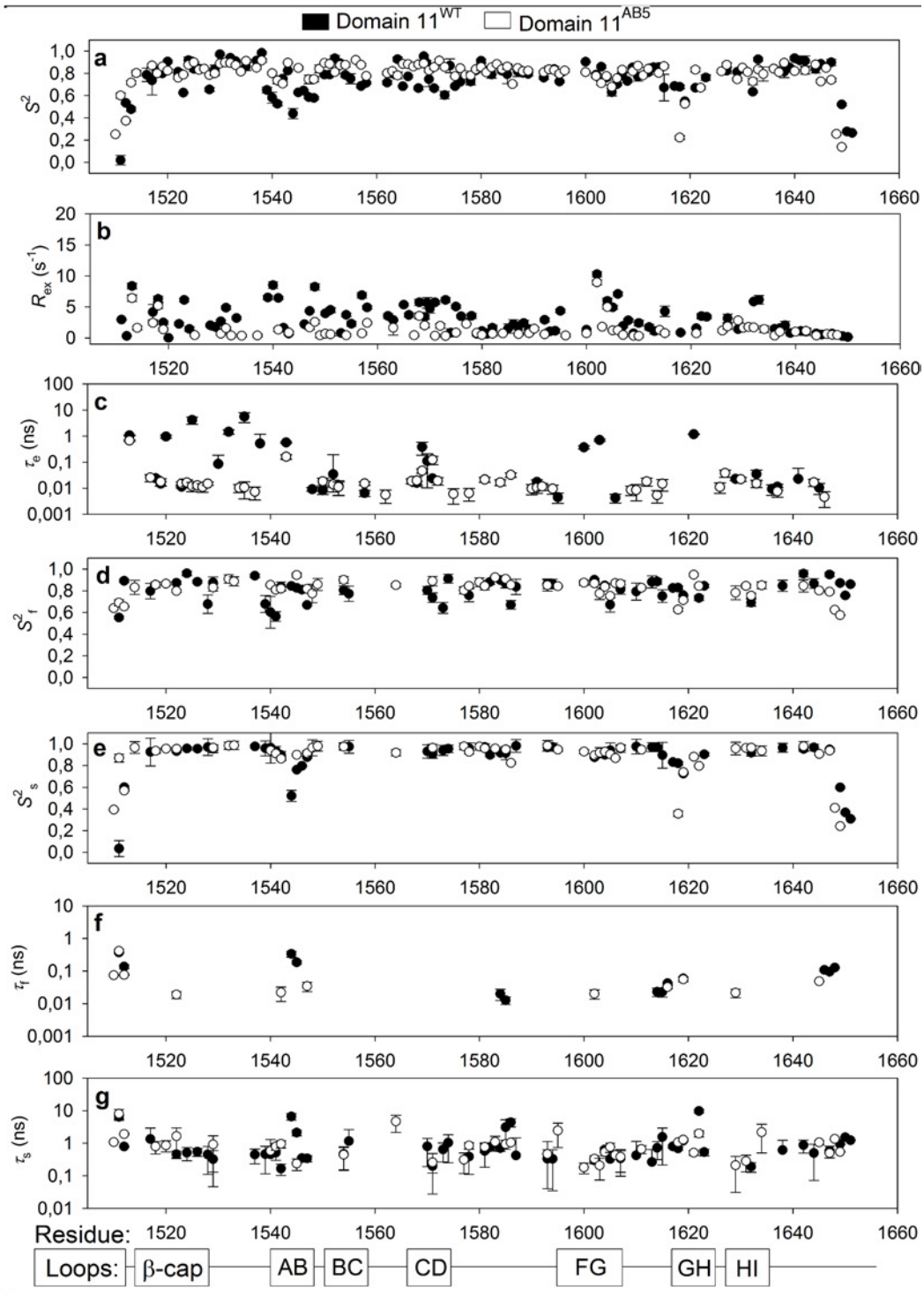
Supplementary Figure S3.

***Fη*, F_{R2} and *J*(0) validation scores for domain 11^{WT} and domain 11^{AB5}** . *Fη*, F_{R2} and *J*(0) validation graphs for domain 11^{WT} (**a-f**) and domain 11^{AB5} (**g-l**). Rows 1 and 3 show individual results for each test where for each residue, the result at 600 (x axis) and 900 (y axis) MHz is correlated, with most consistent residues along the x=y line. Rows 2 and 4 show distribution of data with average value (closer to 1 is more consistent) and standard deviation.

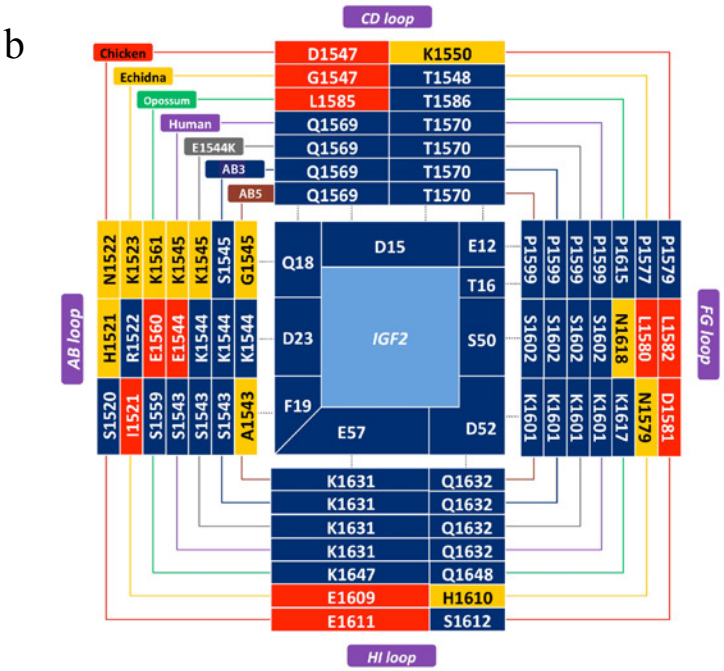


Supplementary Figure S4.

Lipari-Szabo model-free parameters for domain 11^{WT} and domain 11^{AB5}. Graphs showing (a) S^2, (b) R_{ex} and (c) τ_e Lipari-Szabo parameters for domain 11^{WT} (black) and domain 11^{AB5} (white) domain 11 are plotted with respect to residue number. The parameters derived from a two-time scale model (34), (d) S^2_f, (e) S^2_s, (f) τ_f, and (g) τ_s were also fitted Flexible regions have S^2 values lower than 0.7.

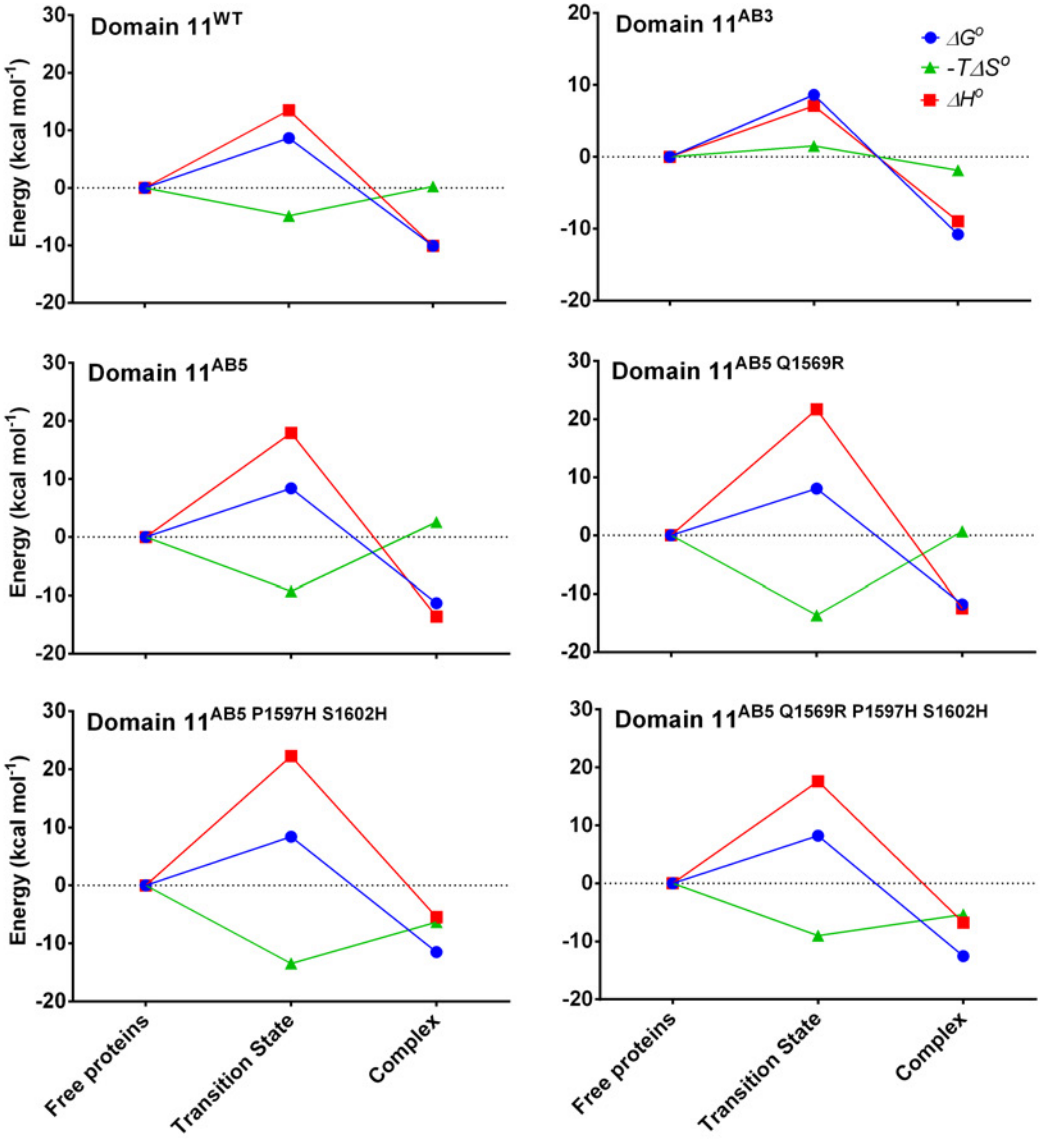


Supplementary Figure S5.
Interaction maps showing the evolution of domain 11 IGF2 binding site residues. (**a**) Interaction map showing the evolution of hydrophobic binding residues on IGF2 (centre) and binding partners (dark blue, indicating favorable hydrophobic interactions) over the AB, CD, FG and HI loops (clockwise). Incorrect charge or polar complementarity is shown in red. Yellow represents where either the acquisition of a charge, or change in steric bulk of a residue, cannot be assessed in the absence of a high-resolution structure of the full complex. (**b**) Complementarity of the non-hydrophobic groups (charged, polar and hydrogen bond interactions) of IGF2 and domain 11 mutants (colors as in **a**.)



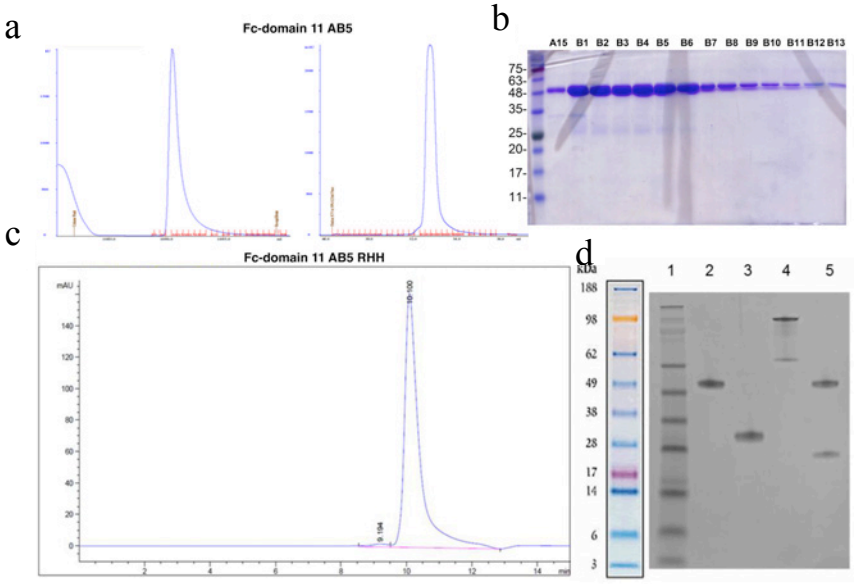
Supplementary Figure S6.

Comparison of thermodynamic parameters for domain 11 mutants and combinations. Surface plasmon resonance was performed with domain 11 mutants and immobilised IGF2 over a range of temperatures (10 °C to 30 °C), van't Hoff and Erying plots generated and *ΔH*, *-TΔS* and *ΔG* values determined. Plots of *ΔH*, *-TΔS* and *ΔG* for free proteins (A+B), transition state (A-B) and complex (AB) for indicated mutants.



Supplementary Figure S7.

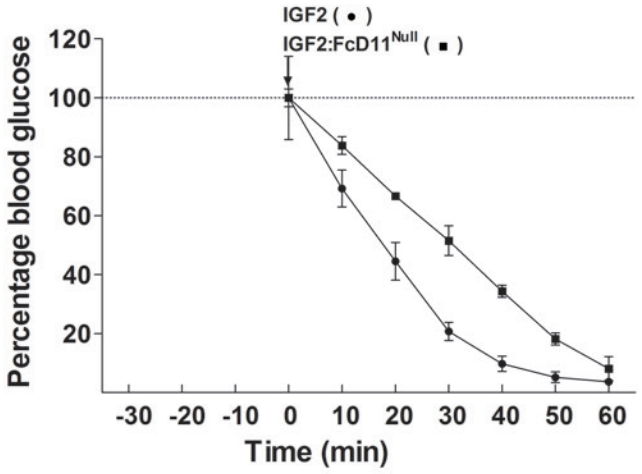
Fc-domain 11^{AB5} and Fc-domain 11 $^{AB5\text{-}RHH}$ fusion protein expression and purification



(a) HEK293T cells were grown in DMEM (10 % FCS and Penicillin/Streptomycin) in a 10L CellStack, transfected with polyethyleneimine (PEI) stock (1 mg/ml, pH 8.0; branched 25kDa PEI, Sigma-Aldrich) with gentle mixing. Following incubation for 4 days, cell supernatant was harvested, clarified by centrifugation (8000g for 30min at 4 °C) and filtration through a 0.45 µM membrane (type HAWP, Millipore) using a vacuum pump and bottle-top filtration unit (Nalgene). Purification was in two stages; initial capture by affinity chromatography followed by ion exchange. Capture was by HiTrap rProtein A FF column (GE Healthcare) pre-equilibrated with 20 mM sodium phosphate pH 7.0 and elution was with 0.1 M sodium citrate pH 3.5 with neutralisation by 0.1 volumes of 1 M Tris pH 9.0. Cation exchange was with a MonoS 5/50GL column (GE Healthcare) pre-equilibrated with 50 mM sodium phosphate pH 7.4 and elution by steep sodium chloride gradient (left panel). Size Exclusion Chromatography was performed with a Superose 6 10/300GL analytical column, pre-equilibrated with 150 mM sodium chloride 50 mM sodium phosphate pH 7.0 (right panel). (b) Purity and specificity were assessed by SDS-PAGE followed by Coomassie staining or western blotting utilising HRP-conjugated anti-Fc antibody (goat polyclonal; Sigma-Aldrich) or HRP-conjugated anti-His tag antibody (Clone BMG-His-1; Roche) as probes. (c) Stable CHOK1SV GS-KO cells were cultured (Lonza proprietary media, 4-10L) in a shaking incubator (140rpm) as described in material and methods. After 7 days, clarified supernatant was purified using a HiTrap MabSelect SuRe column (GE Healthcare), washed and protein eluted with 10mM formate, pH 3.5, and neutrolised with sodium hydroxide. SE-HPLC of the protein (shown) was with a Zorbax GF-250 column (Agilent).(d) Western blot analysis of the purified Fc-domain 11$^{AB5\ RHH}$ protein (lane 2, non reduced, lane 3 reduced) and IgG controls are shown (lane 4, IgG control non reduced, lane 5, IgG control reduced).

Supplementary Figure S8.

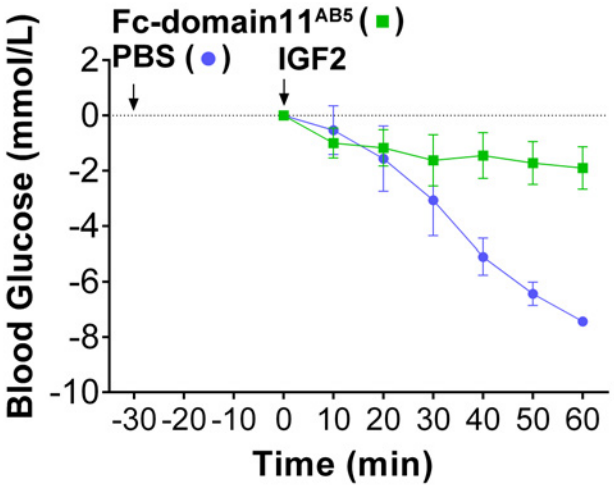
***In vivo* effect of Fc-domain 11 $^{I1572A\ (null)}$ (control) as an IGF2 antagonist in an IGF2-induced hypoglycaemia mouse model**. Following anesthesia and normalisation of blood glucose levels, either 1 mg kg^{-1} IGF2 alone (circles, n=4), or 1 mg kg^{-1} IGF2 premixed with Fc-domain $11^{I1572A\ (null)}$ (1:1) (squares, n=3) were injected and blood glucose levels monitored. Premixing human IGF2 with Fc-domain 11^{I1572A} does not prevent IGF2-induced hypoglycaemia. 2 way ANOVA with Bonferonni post test, not significant.



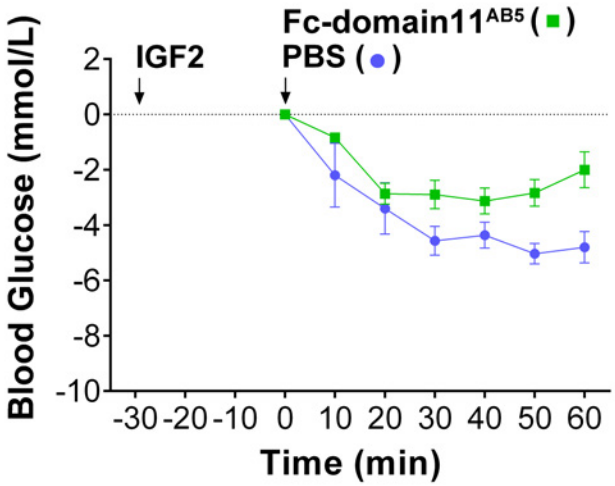
Supplementary Figure S9.

Fc-domain 11^{AB5} abrogates an $IGF2^{1-67}$-induced hypoglycaemia in a mouse model. (**a**) Fc-domain 11^{AB5} (n=4) or PBS (n=3) were injected 30 mins prior to 1 mg kg^{-1} $IGF2^{1-67}$. Preloading of Fc-domain 11^{AB5} abolished the blood glucose decrease caused by subsequent administration of IGF2 (p=0.0358). (b) 1 mg kg^{-1} $IGF2^{1-67}$ was pre-loaded and after 30 min either PBS (n=3) or Fc-domain 11^{AB5} (n=3) were administered, with the later reverting the blood glucose levels towards more positive values. Statistics 2 way ANOVA with Bonferonni post test.

a

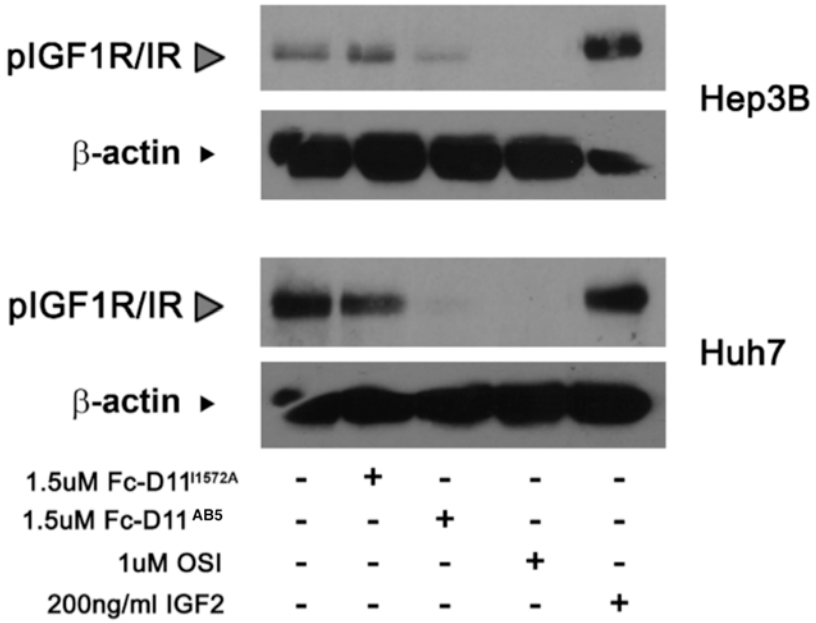


b



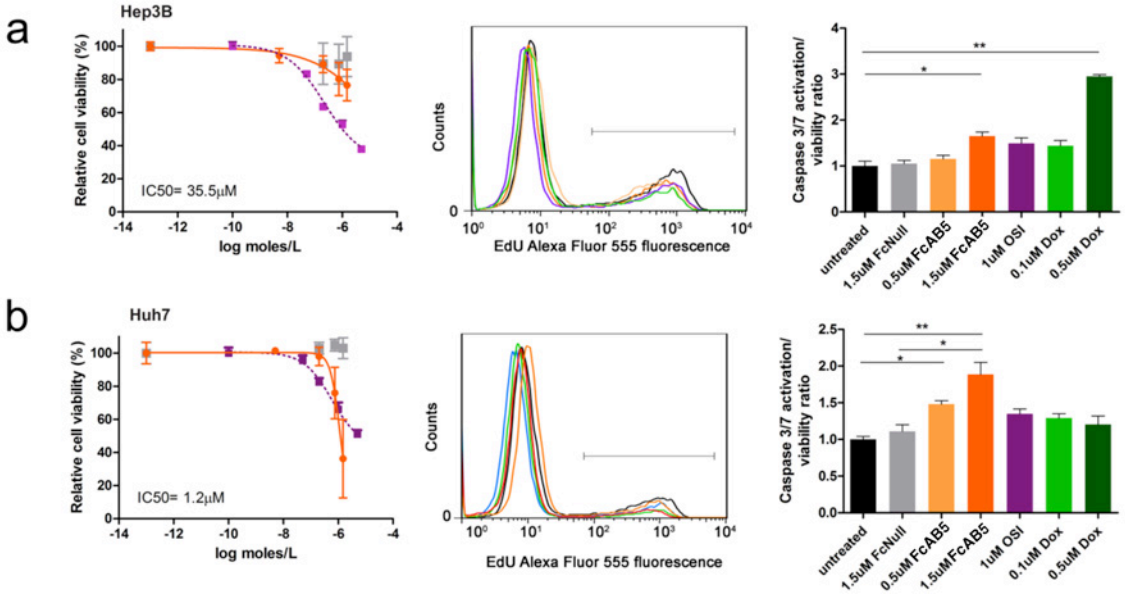
Supplementary Figure S10.

Effect of Fc-domain 11^{AB5} on cell signaling in hepatocellular carcinoma cell lines. The human hepatocellular carcinoma cell lines Hep 3B2.1-7 and Huh-7D12 were obtained from the ATCC (HB-8065) and the ECACC (Catalogue number 01042712), respectively. Hep 3B2.1-7 was grown in EMEM (Lonza) supplemented with 10% (v/v) fetal bovine serum, 100 U mL^{-1} penicillin, 100 µg mL^{-1} streptomycin, and 2mM L-glutamine Huh-7D12 was grown in DMEM (Sigma) supplemented with 10% fetal bovine serum, 100 U mL^{-1} penicillin, 100 µg mL^{-1} streptomycin, and 2mM L-glutamine. The western blot analysis of the effect on the IGF2-TRAP on the IGF2 signaling pathway of both cell lines shows that treatment with the IGF2-TRAP inhibits IGF2-dependent phosphorylation IGF1R/IR-A. Fc-domain 11^{I1572A} and OSI-906 were used as negative and positive control, respectively.



Supplementary Figure S11.

Effect of Fc-domain 11^{AB5} on cell viability, proliferation and apoptosis in hepatocellular carcinoma cell lines.



***In vitro* cell viability.** Cells were plated in 96-well plates at a density of 1-2 x10^4 cells/well in serum free medium. After 24 hr, Fc-domain 11^{AB5} (FcAB5, orange), Fc-domain 11^{I1572A} (FcNull, grey) as negative control or OSI-906 (purple) were added to each well to different final concentrations and the cells were incubated for further 48-72 hr. Cell viability was assayed at 48 and 72 hr after Fc-domain 11^{AB5} treatment using the CellTiter 96® AQueous Non-Radioactive Cell Proliferation Assay (Promega). The half maximal inhibitory concentration (IC_{50}) was calculated using graph prism.

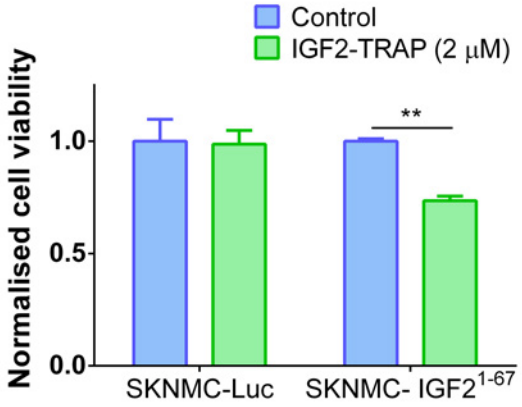
Cell proliferation assay. EdU incorporation was used to measure cell proliferation (Invitrogen, C10338). Briefly, cells were plated at a density of 2x10^5 cells/well in 6-well plates and cultured in serum free media. After 24 hr of serum starvation, PBS (black) and 1.5 µM Fc-domain 11^{I1572} (grey) as negative controls, 0.5 and 1.5 µM Fc-domain 11^{AB5} (orange), 1 µM OSI-906 (purple) and 0.1 and 0.5 µM doxocycline (green) as positive controls were added to the different wells and cells were further cultured. After 48 hr, 10 µM EdU was added for 3 hr and cells were then washed and trypsinised, followed by fixation in 4% PFA and permeabilisation in TBS/5% Triton. The Alexa 555 dye was then attached to the EdU molecules by click chemistry following the manufacturer's instructions and the % of EdU incorporation under different conditions was then analysed by flow cytometry.

Apoptosis assay. Cells were plated in 96-well plates at a density of 1-2 x10^4 cells/well in serum free medium. After 24 hr, PBS (black) and 1.5 µM Fc-domain 11^{I1572} (grey) as negative controls, 0.5 and 1.5 µM Fc-domain 11^{AB5} (orange), 1 µM OSI-906 (purple) and 0.1 and 0.5 µM doxocycline (green) as positive controls were added to different wells. Caspase 3 and 7 activation was assayed at 48 hr after Fc-domain 11^{AB5} treatment using the Caspase-Glo®3/7 Assay (Promega).

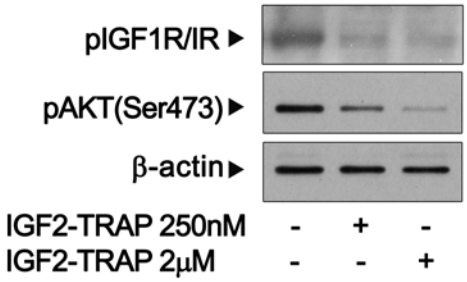
Supplementary Figure S12.

Effect of the IGF2-TRAP on cell signaling and viability of SKNMC-IGF2$^{1-67}$. (**a**) IGF2-TRAP decreases the cell viability of IGF2-autocrine SKNMC-IGF2$^{1-67}$ cell line compared to luciferase only SKNMC control. (**b**) Western blot analysis of the effect on the IGF2-TRAP on the IGF2 signaling pathway of the SKNMC-IGF2$^{1-67}$ cell line shows that treatment with the IGF2-TRAP inhibits IGF2-dependent phosphorylation IGF1R/IR-A and AKT.

a

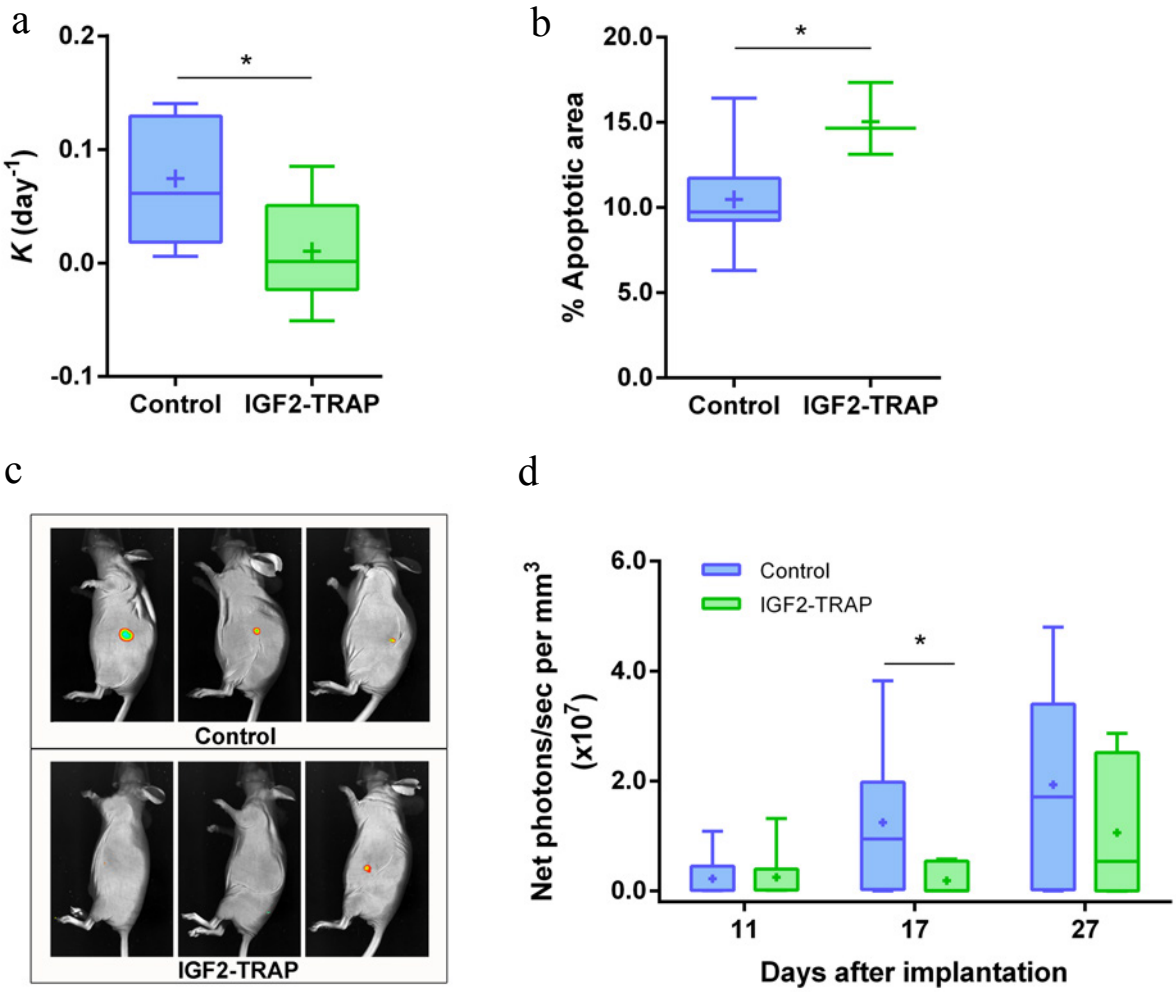


b



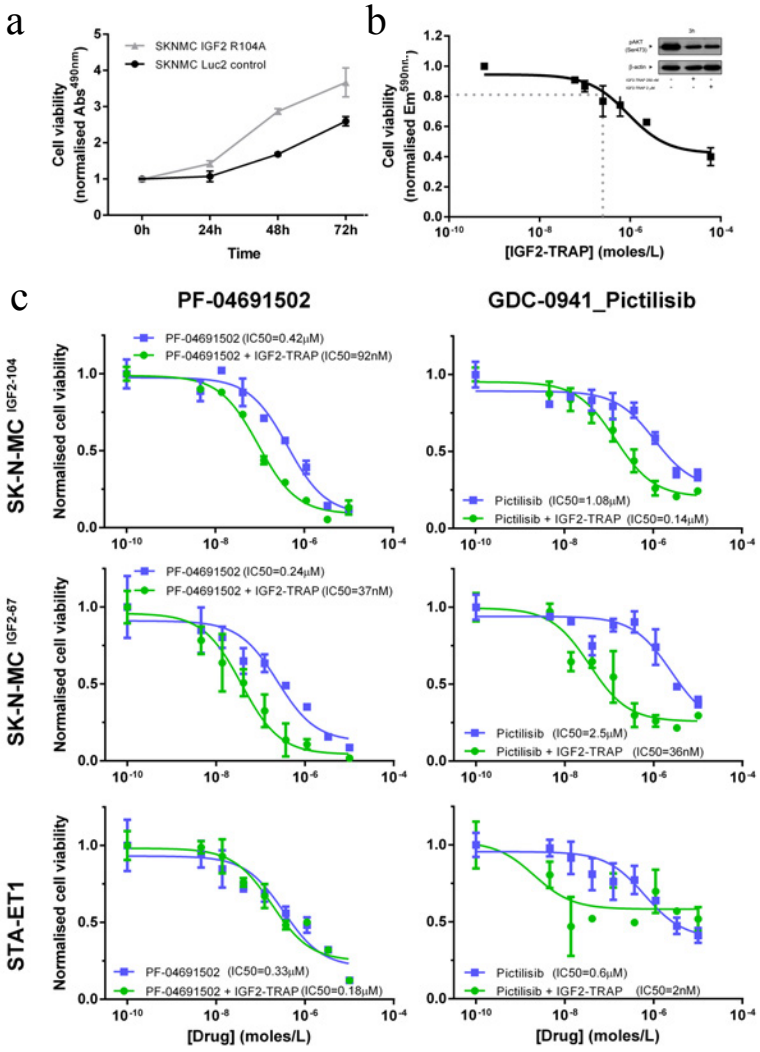
Supplementary Figure S13

Tumour xenografts of Ewing sarcoma cell lines expressing autocrine IGF2 (SKNMC-lucifierase-IGF2^{67}) exhibit reduce growth rates and necrosis with IGF2-TRAP (Fc-domain 11$^{AB5\text{-}RHH}$). (**a**) Growth rate constant (*K*) of xenografts derived from SKNMC-IGF2$^{1\text{-}67}$ cell line is slower in mice receiving the IGF2-TRAP 24 hrs prior to cell implantation, as a single infused concentration of 40 mg kg^{-1} per week over 4 weeks (p=0.036, Mann Whitney test). (**b**) The area of the apoptosis in tumour sections from day 28 tumours following completion of infusion, analysed by H&E staining, increased significantly in treated mice (p=0.049, t-test). (**c**) Tumour growth was monitored by bioluminescence and (**d**) the ratio of signal per tumour volume shows a significant decrease in treated mice after 17 days (p=0.0415, Mann Whitney test), suggesting increased necrosis.



Supplementary Figure S14

PI3 kinase inhibitors exhibit synergistic interactions with the IGF2-TRAP (Fc-domain 11$^{AB5\text{-}RHH}$) in SKNMC-FLAG-IGF2$^{1\text{-}104}$, SKNMC-IGF2$^{1\text{-}67}$ and STA-ET1. (**a**) Cell viability (growth) assays displaying effects of SKNMC-FLAG-IGF2$^{1\text{-}104}$ autocrine IGF2 expression compared to luciferase only SKNMC controls. (**b**) Effect of the IGF2-TRAP on the cell viability of SKNMC-FLAG-IGF2$^{1\text{-}104}$. A concentration of 250 nM of the IGF2-TRAP shows a 10-20% reduction in cell viability and was chosen as the fixed concentration. Insert: Western blot analysis of the effect on the IGF2-TRAP on the IGF2 signaling pathway of the SKNMC-FLAG-IGF2^{104} cell line shows that treatment with the IGF2-TRAP inhibits phosphorylation of pAKT (Ser473). (**c**) Dose response curves for the selected PI3kinase inhibitors (PF-04691502 and Pictilisib) in the presence (green) and absence (blue) of the IGF2-TRAP. Cell viability (Abs 490nm for MTS assay) is normalised to the untreated cells viability. IC_{50} values are shown. A leftward shift of the dose response curve for the combination of the drugs with IGF2-TRAP when compared with single-agent curves indicates synergism ($P<0.0001$ when comparing the IC_{50} of drug alone vs drug + IGF2-TRAP using the F-test). Pictilisib exhibits synergistic activity with STA-ET1. See Figure 6D and 6E.



References for Supporting Information Appendix

1. Wojciechowicz D, Lu CF, Kurjan J, & Lipke PN (1993) Cell surface anchorage and ligand-binding domains of the Saccharomyces cerevisiae cell adhesion protein alpha-agglutinin, a member of the immunoglobulin superfamily. *Mol Cell Biol* 13(4):2554-2563.
2. Wu S & Letchworth GJ (2004) High efficiency transformation by electroporation of Pichia pastoris pretreated with lithium acetate and dithiothreitol. *Biotechniques* 36(1):152-154.
3. Chao G*, et al.* (2006) Isolating and engineering human antibodies using yeast surface display. *Nat Protoc* 1(2):755-768.
4. Zaccheo OJ*, et al.* (2006) Kinetics of insulin-like growth factor II (IGF-II) interaction with domain 11 of the human IGF-II/mannose 6-phosphate receptor: function of CD and AB loop solvent-exposed residues. *J Mol Biol* 359(2):403-421.
5. Williams C*, et al.* (2012) An exon splice enhancer primes IGF2:IGF2R binding site structure and function evolution. *Science* 338(6111):1209-1213.
6. Lescop E, Schanda P, Rasia R, & Brutscher B (2007) Automated spectral compression for fast multidimensional NMR and increased time resolution in real-time NMR spectroscopy. *J Am Chem Soc* 129(10):2756-2757.
7. Delaglio F*, et al.* (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277-293.
8. Vranken WF*, et al.* (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59(4):687-696.
9. d'Auvergne EJ & Gooley PR (2008) Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor. *J Biomol NMR* 40(2):121-133.
10. d'Auvergne EJ & Gooley PR (2008) Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *J Biomol NMR* 40(2):107-119.
11. Clore GM*, et al.* (1990) Deviations from the Simple 2-Parameter Model-Free Approach to the Interpretation of N-15 Nuclear Magnetic-Relaxation of Proteins. *Journal of the American Chemical Society* 112(12):4989-4991.
12. Lipari G & Szabo A (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic- Resonance Relaxation in Macromolecules .1. Theory and Range of Validity. *Journal of the American Chemical Society* 104(17):4546-4559.
13. Lipari G & Szabo A (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic- Resonance Relaxation in Macromolecules .2. Analysis of Experimental Results. *Journal of the American Chemical Society* 104(17):4559-4570.
14. d'Auvergne EJ & Gooley PR (2003) The use of model selection in the model-free analysis of protein dynamics. *J Biomol NMR* 25(1):25-39.
15. d'Auvergne EJ & Gooley PR (2006) Model-free model elimination: a new step in the model-free dynamic analysis of NMR relaxation data. *J Biomol NMR* 35(2):117-135.
16. Prince SN, Foulstone EJ, Zaccheo OJ, Williams C, & Hassan AB (2007) Functional evaluation of novel soluble insulin-like growth factor (IGF)-II-

specific ligand traps based on modified domain 11 of the human IGF2 receptor. *Mol Cancer Ther* 6(2):607-617.
17. Breitling R, Armengaud P, Amtmann A, & Herzyk P (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* 573(1-3):83-92.
18. Ling C*, et al.* (2012) Enhanced radiosensitivity of non-small-cell lung cancer (NSCLC) by adenovirus-mediated ING4 gene therapy. *Cancer gene therapy* 19(10):697-706.
19. Rieping W*, et al.* (2007) ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* 23(3):381-382.
20. Doreleijers JF*, et al.* (2012) NRG-CING: integrated validation reports of remediated experimental biomolecular NMR data and coordinates in wwPDB. *Nucleic Acids Res* 40(Database issue):D519-524.
21. Winn MD*, et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235-242.
22. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* 63(Pt 1):32-41.
23. Murshudov GN*, et al.* (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355-367.
24. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
25. Laskowski RA, Moss DS, & Thornton JM (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* 231(4):1049-1067.
26. Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12-21.
27. Joosten RP*, et al.* (2009) PDB_REDO: automated re-refinement of X-ray structure models in the PDB. *J Appl Crystallogr* 42(Pt 3):376-384.
28. Brown J*, et al.* (2008) Structure and functional analysis of the IGF-II/IGF2R interaction. *EMBO J* 27(1):265-276.
29. Fushman D, Tjandra N, & Cowburn D (1998) Direct measurement of N-15 chemical shift anisotropy in solution. *J. Am. Chem. Soc.* 120(42):10947-10952.
30. Farrow NA, Zhang OW, Szabo A, Torchia DA, & Kay LE (1995) Spectral Density-Function Mapping Using N-15 Relaxation Data Exclusively. *Journal of Biomolecular Nmr* 6(2):153-162.
31. Morin S & Gagne SM (2009) Simple tests for the validation of multiple field spin relaxation data. *Journal of Biomolecular Nmr* 45(4):361-372.
32. Fushman D, Tjandra N, & Cowburn D (1999) An approach to direct determination of protein dynamics from N-15 NMR relaxation at multiple fields, independent of variable N-15 chemical shift anisotropy and chemical exchange contributions. *J. Am. Chem. Soc.* 121(37):8577-8582.
33. Miletti T, Farber PJ, & Mittermaier A (2011) Active site dynamics in NADH oxidase from Thermus thermophilus studied by NMR spin relaxation. *Journal of Biomolecular Nmr* 51(1-2):71-82.
34. Clore GM*, et al.* (1990) Deviations from the Simple 2-Parameter Model-Free Approach to the Interpretation of N-15 Nuclear Magnetic-Relaxation of Proteins. *J. Am. Chem. Soc.* 112(12):4989-4991.